

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	2327	89.1	496	2	Q96DK0	Q96dk0	homo sapien
2	2295.5	87.9	500	2	Q9BRV0	Q9brv0	homo sapien
3	2262.5	86.7	497	2	Q8WY24	Q8wy24	homo sapien
4	2224.5	85.2	499	2	Q8NSK4	Q8nsk4	homo sapien
5	2215	84.8	494	2	Q96K68	Q96k68	homo sapien
6	2211	84.7	506	2	Q6N090	Q6n090	homo sapien
7	2209	84.6	506	2	Q6MZW0	Q6mzw0	homo sapien
8	2204.5	84.4	493	2	Q6GCMX2	Q6gmx2	homo sapien
9	2201.5	84.3	493	2	Q8NCL6	Q8nc16	homo sapien
10	2196.5	84.1	519	2	Q6N092	Q6n092	homo sapien
11	2191	83.9	494	2	Q6ZMW64	Q6zwm64	homo sapien
12	2131	81.6	496	2	Q96KX8	Q96kx8	homo sapien
13	2103.5	80.6	480	2	Q6P089	Q6p089	homo sapien
14	2088	80.0	498	2	Q6N041	Q6n041	homo sapien
15	2087.5	80.0	500	2	Q6N091	Q6n091	homo sapien
16	2032.5	77.8	479	2	Q6MZV6	Q6mzv6	homo sapien
17	2017.5	77.3	483	2	Q6MZV9	Q6mzv9	homo sapien
18	2012.5	77.1	487	2	Q6ZVX0	Q6zvx0	homo sapien
19	1982.5	75.9	477	2	Q6GCMX7	Q6gmx7	homo sapien
20	1969	75.4	384	2	Q9UP60	Q9up60	homo sapien
21	1962.5	75.2	478	2	Q72379	Q7z379	homo sapien
22	1947	74.6	492	2	Q72374	Q7z374	homo sapien
23	1946	74.5	478	2	Q6NHY3	Q6nyh3	homo sapien
24	1899	72.7	416	2	Q9NPP6	Q9npp6	homo sapien
25	1889	72.3	353	1	ALC1_HUMAN	P01876	homo sapien
26	1842	70.5	353	1	ALC1_GORGO	P20758	gorilla gorilla
27	1710.5	65.5	340	1	ALC2_HUMAN	P01877	homo sapien
28	1455	55.7	481	2	Q91WT1	Q91wt1	mus musculus
29	1449.5	55.5	482	2	Q8K172	Q8k172	mus musculus
30	1442.5	55.2	488	2	Q91WR1	Q91wr1	mus musculus
31	1439.5	55.1	489	2	Q8VCX4	Q8vcx4	mus musculus





```

RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi Y., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Negahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta K., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Nomiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Taksumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs."
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK027379; BAB55072.1; -.
DR PIR; S21205; S21205.
DR HSSP; P01878; ICOW.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 2.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 494 AA; 53088 MW; 9ALDTAE5ABEA4C0E CRC64;

Query Match 84.8%; Score 2215; DB 2; Length 494;
Best Local Similarity 84.8%; Pred. No. 4.1e-125;
Matches 417; Conservative 26; Mismatches 41; Indels 8; Gaps 2;

QY 5 GFLWALVISTCLFBSMAQVLQVSGAEVKKPGSSVRVSKASGGTFSYVAISWVRQAPG 64
DB 4 GLRWVFLVA-FLBGVQCEVLVSGGLVKPGGSLRLSCAASGLSTYAMNWRQAPG 62

QY 65 GLEWMGGIIPFTANYAQKFGQGRVTITADESTSTAYMELSSLRSDTAVYYCARDPF-- 122
DB 63 GLEWSSISRSRDIYIYRDSVKGKRTISRDNKNSLYLQWNSLRVDVTAVYYCARDSCNG 122

QY 123 -----LHYWGQGLTVTVSTASPTSPKVPFLSLCSTQPDGNNVIACTLVQGFPPPLSVTM 177
DB 123 AICYGSPGQGLTVTVSSASPTSPKVPFLSLCSTQPDGNNVIACTLVQGFPPPLSVTM 182

QY 178 SESGQGVTAARNFPFSDASGDLTYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVP 237
DB 183 SESGQGVTAARNFPFSDASGDLTYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVP 242

QY 238 PVPSTPTPTSPPTPTSPSCCHPRLSLRPALEDLLGSEANLTCTLTGLRDASGVTTT 297
DB 243 PVPSTPTPTSPPTPTSPSCCHPRLSLRPALEDLLGSEANLTCTLTGLRDASGVTTT 302

QY 298 WTPSSGKSAVQGPDRDLGCGYSSVSLGCAEPNHNKGTFTCTAAYPSKPTLTATLSK 357
DB 303 WTPSSGKSAVQGPDRDLGCGYSSVSLGCAEPNHNKGTFTCTAAYPSKPTLTATLSK 362

```

## RESULT 6

Q6N090

ID Q6N090 PRELIMINARY; PRT; 506 AA.

AC Q6N090; DT 05-JUL-2004 (TEMBLrel. 27, Created)

DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)

DE Hypothetical protein DKFp686G21220 (Fragment).

GN Names=DKFp686G21220;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]\_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=Human rectum tumor;

RG The German Human cDNA Consortium;

RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,

RA Fobo G., Han M., Wiemann S.;

RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BX640626; CAE45780.1; -.

DR HSSP; P01751; IA6W.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig\_c1.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003596; Ig\_v.

DR Pfam; PF07654; C1-set; 2.

DR SMART; SM00409; IG; 4.

DR SMART; SM00407; IGc1; 2.

DR SMART; SM00406; IGV; 1.

DR PROSITE; PS00835; IG\_LIKE; 4.

DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.

KW Hypothetical protein.

SQ SEQUENCE 506 AA; 54388 MW; CCC670F063D32612 CRC64;

Query Match 84.7%; Score 2211; DB 2; Length 506;

Best Local Similarity 85.5%; Pred. No. 7.3e-125;

Matches 420; Conservative 17; Mismatches 44; Indels 10; Gaps 3;

QY 5 GFLWA-LVISTCLFBSMAQVLQVSGAEVKKPGSSVRVSKASGGTFSYVAISWVRQAPG 63

DB 20 GWTWSLFLVAATTGAPSOAQLQSGPEVKLFGASVKVCSNPDTDYKFGSYDISWRQAPG 79

QY 64 QGLEWNGGIIPFTANYAQKFGQGRVTITADESTSTAYMELSSLRSDTAVYYCAR---- 119

DB 80 QGLEWNGWIGLVDGNARPTQNFEDRINIITKDTSTNIYVMEMRGLTPDDTATYFCVREGTAV 139

QY 120 -DPFLHYWGQGLTVTVSTASPTSPKVPFLSLCSTQPDGNNVIACTLVQGFPPPLSVTWS 178

DB 140 LDP----WGQGLTVTVSSASPTSPKVPFLSLCSTQPDGNNVIACTLVQGFPPPLSVTWS 195

QY 179 ESGQGVTAARNFPFSDASGDLTYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVP 238

DB 196 ESGQGVTAARNFPFSDASGDLTYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVP 255

QY 239 VPSTPTPTSPPTPTPTSPSCCHPRLSLRPALEDLLGSEANLTCTLTGLRDASGVTTT 298

DB 256 VPSTPTPTSPPTPTPTSPSCCHPRLSLRPALEDLLGSEANLTCTLTGLRDASGVTTT 315

```
Qy 299 TPSSGKSAVQPPDRDLGCGYSSVLSGCAEPNHNHGTFTCTAAAYPESKTPPLTATLSKS 358
Db 316 TPSSGKSAVQPPDRDLGCGYSSVLSGCAEPNHNHGTFTCTAAAYPESKTPPLTATLSKS 375
Qy 359 GNTFRPEVHLLPPPSSEELALNELVTLCLARGFSPKDVLRWLGSGOELPREKYLTVASR 418
Db 376 GNTFRPEVHLLPPPSSEELALNELVTLCLARGFSPKDVLRWLGSGOELPREKYLTVASR 435
Qy 419 QEPSQGGTTTFAVTSILRVAAEDMKKGGDTFSCMVGHEALPLAFTQKTDRLAGKPTHNVNS 478
Db 436 QEPSQGGTTTFAVTSILRVAAEDMKKGGDTFSCMVGHEALPLAFTQKTDRLAGKPTHNVNS 495
Qy 479 VMAEVDGTCY 489
Db 496 VMAEVDGTCY 506

RESULT 7
Q6GMZW0 Q6GMZW0 PRELIMINARY; PRT; 506 AA.
AC Q6GMZW0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DRF2p686j11235 (Fragment).
GN Name=DRF2p686j11235;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human small intestine;
RG The German Human cDNA Consortium;
RA Bloeker H., Boscher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640847; CAE45917.1; -.
DR HSSP; P01751; 1A6W.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGc1; 2.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS00835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
FT NON TRP
SQ SEQUENCE 506 AA; 54459 MW; 1D9477A37B77C3A0 CRC64;

Query Match 84.6%; Score 2209; DB 2; Length 506;
Best Local Similarity 88.3%; Pred. No. 9.7e-125;
Matches 416; Conservative 18; Mismatches 35; Indels 2; Gaps 1;

Qy 21 AQVOLVSGAEVKVPGSSVRVSCASGGTSSYAIISVRQAPGQGLEWGGIIPFGTAN 80
Db 36 SQVOLVSGAEVKVPGASVLSKSCPSGNTVNIYALHLRQAPGHPGEMTWLNVENGMM 95
Qy 81 YAKFOGRTVTADSTSTAYMELSLRSEDTAVYCA--RDPLHYWGQTLTVGTAS 138
Db 96 SSHFRRLTVTRDASANTVYMLTSLTSETAVYFCASERGLGVWVGPTTVTVSSAS 155
Qy 139 PTPSKVPFLSLCSTPPDGNVVIACLVQGFPPQEPFLSVTWSESGQGVTVARNFPPSQDASGD 198
Db 156 PTPSKVPFLSLCSTPPDGNVVIACLVQGFPPQEPFLSVTWSESGQGVTVARNFPPSQDASGD 215
Qy 199 LYTTSSQLTLPATQCLAGKSVTCHVKIYTNPSQDVTVPSPPTPPTPSPSPSC 258
```

```
Db 216 LYTTSSQLTLPATQCLAGKSVTCHVKIYTNPSQDVTVPSPPTPPTPSPSPSC 275
Qy 259 CHPLSLRHPALDELGLLGSEANLTCTLTGLRDASGVTFTTTPSSGKSAVQPPDRDLGCG 318
Db 276 CHPLSLRHPALDELGLLGSEANLTCTLTGLRDASGVTFTTTPSSGKSAVQPPDRDLGCG 335
Qy 319 YSVSSVLSGCAEPNHNHGTFTCTAAAYPESKTPPLTATLSKSGNTFRPEVHLLPPPSSEELAL 378
Db 336 YSVSSVLSGCAEPNHNHGTFTCTAAAYPESKTPPLTATLSKSGNTFRPEVHLLPPPSSEELAL 395
Qy 379 NELVTLTCLARGFSPKDVLRWLGSGOELPREKYLTVASRQEPSQGGTTTFAVTSILRVAA 438
Db 396 NELVTLTCLARGFSPKDVLRWLGSGOELPREKYLTVASRQEPSQGGTTTFAVTSILRVAA 455
Qy 439 EDWKKGDTFSCMVGHEALPLAFTQKTDRLAGKPTHNVNSVMAEVDGTCY 489
Db 456 EDWKKGDTFSCMVGHEALPLAFTQKTDRLAGKPTHNVNSVMAEVDGTCY 506

RESULT 8
Q6GMX2 Q6GMX2 PRELIMINARY; PRT; 493 AA.
AC Q6GMX2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22398257; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieff F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC073771; AAH73771.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 2.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGc1; 2.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS00835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
```

```
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 52865 MW; 55B999305B286203 CRC64;

Query Match      84.4%; Score 2204.5; DB 2; Length 493;
Best Local Similarity 84.5%; Pred. No. 1.9e-124;
Matches 415; Conservative 26; Mismatches 43; Indels 7; Gaps 2;

QY 5 GFLWALVISTCLFESMAQVQLVQSGAEVKKPGSSVRVSKASGGTFSSVAISWVRAPQG 64
DB 4 GLSWVFLVA-ILKGVCQEVQLVSGGVLQVPGSLRLSCAASGFISSVWMHVRAPGK 62
QY 65 GLEWGGIIPFGTAN-VAQKQGVITITADESTSTAYMELSSLRSESDTAVVYCARD- 120
DB 63 GLVWVSRKNSDGSSTVADSVKGRFTISRDNKNTLYLQWNSLRGDEAAVYVCARGFVSL 122
QY 121 --PFLHYWGQGLTVTVSTASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWS 178
DB 123 PRSTLDIWGQGLTVTVSTASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWS 182
QY 179 ESGQGVTVARNFPSPDASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCP 238
DB 183 ESGQGVTVARNFPSPDASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCP 242
QY 239 VPSTPTPTSPSTPTPTSPSCCHPRLSLHRPALEDLLLGSEANLTCITLTGLRDASGVTF 298
DB 243 VPSTPTPTSPSTPTPTSPSCCHPRLSLHRPALEDLLLGSEANLTCITLTGLRDASGVTF 302
QY 299 TPSSGKSAVQGPDPDRDLCCGYSVSVSLPGCAEPWNHGKFTCTAAYPESKTPLTATLSK 358
DB 303 TPSSGKSAVQGPDPDRDLCCGYSVSVSLPGCAEPWNHGKFTCTAAYPESKTPLTATLSK 362
QY 359 GNTFRPEVHLLPPPSSEELALNELVLTCLARGSPKDVLRVWLQSGOELPREKYLTVASR 418
DB 363 GNTFRPEVHLLPPPSSEELALNELVLTCLARGSPKDVLRVWLQSGOELPREKYLTVASR 422
QY 419 QEPSQGTTFITFAVTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNS 478
DB 423 QEPSQGTTFITFAVTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNS 482
QY 479 VVMAEVDGTCY 489
DB 483 VVMAEVDGTCY 493

RESULT 9
Q8NCL6 PRELIMINARY; PRT; 493 AA.
AC Q8NCL6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ90170.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishi S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074651; BAC11114.1; -
DR HSSP; P01876; IOW0.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 2.
DR SMART; SM00406; IGV; 1.
```

```
DR PROSITE; PS50835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
SQ SEQUENCE 493 AA; 53224 MW; 12ECD7E094777101 CRC64;

Query Match      84.3%; Score 2201.5; DB 2; Length 493;
Best Local Similarity 84.8%; Pred. No. 2.7e-124;
Matches 417; Conservative 23; Mismatches 43; Indels 9; Gaps 4;

QY 5 GFLWALVISTCLFESMAQVQLVQSGAEVKKPGSSVRVSKASGGTFSSVAISWVRAPQG 64
DB 4 GLSWVFLVA-LLRGVQCQVQLVESGGVLPGLSLRLSCAASGFRPRDMDHWRVRSQPE 62
QY 65 GLEWGGIIPFGTAN-VAQKQGVITITADESTSTAYMELSSLRSESDTAVVYCARDP-- 121
DB 63 GLEWV-ALIWYDGTXTYISDSVKGRLTVSRDNYKNTLYLHMKSLGAEDTAVVYCARDQGY 121
QY 122 ----FLHYWGQGLTVTVSTASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWS 177
DB 122 AGVGVFDHMGQGLTVTVSTASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWS 181
QY 178 ESGQGVTVARNFPSPDASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCP 237
DB 182 ESGQGVTVARNFPSPDASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCP 241
QY 238 VPSTPTPTSPSTPTPTSPSCCHPRLSLHRPALEDLLLGSEANLTCITLTGLRDASGVTF 297
DB 242 VPSTPTPTSPSTPTPTSPSCCHPRLSLHRPALEDLLLGSEANLTCITLTGLRDASGVTF 301
QY 298 WTPSSGKSAVQGPDPDRDLCCGYSVSVSLPGCAEPWNHGKFTCTAAYPESKTPLTATLSK 357
DB 302 WTPSSGKSAVQGPDPDRDLCCGYSVSVSLPGCAEPWNHGKFTCTAAYPESKTPLTATLSK 361
QY 358 GNTFRPEVHLLPPPSSEELALNELVLTCLARGSPKDVLRVWLQSGOELPREKYLTVAS 417
DB 362 GNTFRPEVHLLPPPSSEELALNELVLTCLARGSPKDVLRVWLQSGOELPREKYLTVAS 421
QY 418 ROEPSQGTTFITFAVTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNS 477
DB 422 ROEPSQGTTFITFAVTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNS 481
QY 478 VVMAEVDGTCY 489
DB 482 VVMAEVDGTCY 493

RESULT 10
Q6N092 PRELIMINARY; PRT; 519 AA.
AC Q6N092;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K18196;
GN Name=DKFZp686K18196;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human esophagus tumor;
RG The German Human cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640624; CAE45778.1; -
DR HSSP; P01842; 1AOK.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 2.
```

[illegible]

```
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Glandular pool- thyroid;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Glandular pool- thyroid;
RX Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RA EMBL; BC065733; AAH65733.1; -.
DR HSSP; P01751; 1A6W.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig.MHC.
DR InterPro; IPR003596; Ig.v.
DR Pfam; PF07654; C1-set; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; Igv; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
DR Hypothetical protein.
SQ SEQUENCE 480 AA; 51997 MW; 2E286C57E4F0ED65 CRC64;
```

Query Match 81.6%; Score 2131; DB 2; Length 496;

Best Local Similarity 81.8%; Pred. No. 4.6e-120;

Matches 404; Conservative 29; Mismatches 49; Indels 12; Gaps 4;

7 LW-ALVISTCLFSPMAQVQVLOSAGVKKPGSSVRVSKASGGTF--SSYATISWVRQAPG 63

4 LKFFLLVLAAPRWLSQLQESGPGVLKSSSTLSLTCVSGSSISSSSYNGWNRQPG 63

64 QGLWMGGIIPITGTANYAQFGGRVTITADSTSTAYMELSLRSEDVAVYCARDPF- 122

64 KGLEWIANITY-YSGITYYNPSSLKSRVTISVDTSKNQLSLKVRSVTAADTAVYFCARHGY 122

123 -----LHYWGQGLVTVSTASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQPLSV 175

123 RSGRTGAIDYWGQGLTVTVSSASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQPLSV 182

176 TWSESGQGTARNFPSPQASGDLVTTSSQLTLPATQCLAGKSVTCVHKYTNPSQDVTV 235

183 TWSESGQGTARNFPSPQASGDLVTTSSQLTLPATQCLAGKSVTCVHKYTNPSQDVTV 242

236 PCVPPT 295

243 PCVPPT 302

296 FTWTPSSGKSAVGGPDRDLCGCVSVSVLSGCAEPNHNKFTCTAAVPESKTPLTATL 355

303 FTWTPSSGKSAVGGPDRDLCGCVSVSVLSGCAEPNHNKFTCTAAVPESKTPLTATL 362

356 SKSGNTFRPEVHLPPPSBELALNELVTLTCLARGSPKDLVLRWLQSGQELPREKYLFW 415

363 SKSGNTFRPEVHLPPPSBELALNELVTLTCLARGSPKDLVLRWLQSGQELPREKYLFW 422

416 ASRQEPSQGTPTTFAVTSILRVAADWKKGDTFSCWVGHEALPLAFTQKTDILAGKPTHV 475

423 ASRQEPSQGTPTTFAVTSILRVAADWKKGDTFSCWVGHEALPLAFTQKTDILAGKPTHV 482

OY 476 NVSVVMAEVDGTCY 489

DB 483 NVSVVMAEVDGTCY 496

#### RESULT 13

Q6P089 PRELIMINARY; PRT; 480 AA.

AC Q6P089;

DT 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE Hypothetical protein.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Glandular pool- thyroid;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Glandular pool- thyroid;

RA Strausberg R.;

RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC065733; AAH65733.1; -.

DR HSSP; P01751; 1A6W.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig-cl.

DR InterPro; IPR003006; Ig.MHC.

DR InterPro; IPR003596; Ig.v.

DR Pfam; PF07654; C1-set; 2.

DR SMART; SM00409; IG; 4.

DR SMART; SM00407; IGcl; 3.

DR SMART; SM00406; Igv; 1.

DR PROSITE; PS50835; IG\_LIKE; 4.

DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_2.

DR Hypothetical protein.

SQ SEQUENCE 480 AA; 51997 MW; 2E286C57E4F0ED65 CRC64;

Query Match 80.6%; Score 2103.5; DB 2; Length 480;

Best Local Similarity 84.6%; Pred. No. 2e-118;

Matches 402; Conservative 19; Mismatches 35; Indels 19; Gaps 2;

21 AQVLVQSGAEVKKPGSSVRVSKASGGTFSSYATISWVRQAPQGLWMGGIIPITGTAN 80

19 SQVLVQSGAEVKKPGSSVRVSKASGGTFSSYATISWVRQAPQGLWMGGIIPITGTAN 78

81 YAQKFGGRVTITADSTSTAYMELSLRSEDVAVYCAR-----DPLHYWGQGLTVTV 134

```
Db 79 SAEKFGQVRVTITDTSNTAYMELTSLKSDDTALYYCARGHSDWSSYYFDYWGQGLTVTV 138
Qy 135 STASPTSPKVPFPLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESQGVTAARNFPPSQD 194
Db 139 SSASPTSPKVPFPLSLDSTPDQDGNVVIACLVQGFPPQBPPLSVTWSESQGVTAARNFPPSQD 198
Qy 195 ASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCPVPSTPTPSPSTPTP 254
Db 199 ASGDLVTTSSQLTLPATQCPDGKSVTCHVKHYTNPSQDVTPCPVP-----P 245
Qy 255 SPSCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSASVQGPDPDR 314
Db 246 PPCCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGATFTWTPSSGKSASVQGPDPDR 305
Qy 315 LCGCYSSVSVLSCGCAEPWNNHGKFTCTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSE 374
Db 306 LCGCYSSVSVLPGCAQPMNHGFTCTCTAAPHLPKLTPLTANITKSGNTFRPEVHLLPPSE 365
Qy 375 ELALNELVTLTCLARGSPKDVLRVWLGQSOELPREKYLTVASRQEPSQGTTFPAVTSIL 434
Db 366 ELALNELVTLTCLARGSPKDVLRVWLGQSOELPREKYLTVASRQEPSQGTTFPAVTSIL 425
Qy 435 RVAEADWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 489
Db 426 RVAEADWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 480

RESULT 14
Q6N041 PRELIMINARY; PRT; 498 AA.
ID Q6N041 AC Q6N041 PRELIMINARY; PRT; 498 AA.
AC Q6N041 AC Q6N041 PRELIMINARY; PRT; 498 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686O16217 (Fragment).
GN Names=DKFZp686O16217;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human rectum tumor;
RG The German Human cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640710; CAE45829.1; -
DR HSSP; P01751; IA6W.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; Cl-set; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGC1; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00405; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 498 AA; 54125 MW; 40B3208A84E03B46 CRC64;

Query Match 80.0%; Score 2088; DB 2; Length 498;
Best Local Similarity 82.6%; Pred. No. 1.7e-117;
Matches 39; Conservative 27; Mismatches 34; Indels 22; Gaps 2;

Qy 21 AQVLQVSGAEVKKPGSSVRVSVKASGCTFSSVAISVVRQAPGQGLFWMGIIPIFGTAN 80
Db 34 SQVLQVSGADVKKPGASVSVKASGCTFTNFHFHVRQAPGQGLFWMGIIPIFGTAN 93
Qy 81 YAEKFGQVRVTITDTSNTAYMELTSLKSDDTALYYCARGHSDWSSYYFDYWGQGLTVTV 131
```

```
Db 94 YAEKFGQVRVTITDTSNTAYMELTSLKSDDTALYYCARGHSDWSSYYFDYWGQGLTVTV 153
Qy 132 VTVSTASPTSPKVPFPLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESQGVTAARNFPP 191
Db 154 VTVSSASPTSPKVPFPLSLDSTPDQDGNVVIACLVQGFPPQBPPLSVTWSESQGVTAARNFPP 213
Qy 192 SQDASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCPVPSTPTPSPSTPTP 251
Db 214 SQDASGDLVTTSSQLTLPATQCPDGKSVTCHVKHYTNPSQDVTPCPVP----- 262
Qy 252 PTSPSCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSASVQGP 311
Db 263 --PPPCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGATFTWTPSSGKSASVQGP 320
Qy 312 DRDLCCGYSSVSVLSCGCAEPWNNHGKFTCTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPP 371
Db 321 ERDLCCGYSSVSVLPGCAQPMNHGFTCTCTAAPHLPKLTPLTANITKSGNTFRPEVHLLPP 380
Qy 372 PSBELALNELVTLTCLARGSPKDVLRVWLGQSOELPREKYLTVASRQEPSQGTTFPAV 431
Db 381 PSBELALNELVTLTCLARGSPKDVLRVWLGQSOELPREKYLTVASRQEPSQGTTFPAV 440
Qy 432 SILRVAEADWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 489
Db 441 SILRVAEADWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 498

RESULT 15
Q6N091 PRELIMINARY; PRT; 500 AA.
ID Q6N091 AC Q6N091 PRELIMINARY; PRT; 500 AA.
AC Q6N091 AC Q6N091 PRELIMINARY; PRT; 500 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686C02220 (Fragment).
GN Names=DKFZp686C02220;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human rectum tumor;
RG The German Human cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640625; CAE45779.1; -
DR HSSP; P01751; IA6W.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; Cl-set; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGC1; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 500 AA; 54160 MW; 3C423A17D65A41E4 CRC64;

Query Match 80.0%; Score 2087.5; DB 2; Length 500;
Best Local Similarity 82.8%; Pred. No. 1.9e-117;
Matches 39; Conservative 25; Mismatches 36; Indels 21; Gaps 2;

Qy 21 AQVLQVSGAEVKKPGSSVRVSVKASGCTFSSVAISVVRQAPGQGLFWMGIIPIFGTAN 80
Db 37 SQVLQVSGAEVKKPGASVSVKASGCTFSDHSITHLRQAPGQGLFWMGIIPIFGTAN 96
```

Qy	81	YAKFOGRVITITADESTAYMELSLRSED	TAVVYCARD-----PFLHYWGOTLV	132
Db	97	YAQNLQGRVTMTTDTSTAYMELSLRSD	D TAVYYCAKQSYTTIPNDAFHIWGGT	156
Qy	133	TVSTASPTSPKVFPLSLCSTQPDGNV	VIACLVOGFFPQPPLSVTWSESGQV	192
Db	157	TVSSASPTSPKVFPLSLDSTPQDGNV	VVACLVOGFFPQPPLSVTWSESGQV	216
Qy	193	QDASGLYTTSSQLTLPATQCLAGKSV	TCHVKHYTNPSQDVTVPCVPSTPT	252
Db	217	QDASGLYTTSSQLTLPATQCPDGKSV	TCHVKHYTNSSQDVTVPCRP-----	264
Qy	253	TPSPCCHPRLSLHREPALEDLLLGSE	ANLTCTLTGLRDASGVTFTWTPSSG	312
Db	265	-PPPPCCHPRLSLHREPALEDLLLGSE	ANLTCTLTGLRDASGATFTWTPSSG	323
Qy	313	RDLGGCYSVSSVLSGCAEPNHNHKT	FTCTAAYPESKTPLTATLSKSGNT	372
Db	324	RDLGGCYSVSSVLPGCAQSNHGET	FTCTAAHPELKTPLTANITKSGNT	383
Qy	373	SEELANLVLTLTCLARGFSPKDVLR	WLQGSQELPREKYLTVASRQPSQ	432
Db	384	SEELANLVLTLTCLARGFSPKDVLR	WLQGSQELPREKYLTVASRQPSQ	443
Qy	433	ILRVAEDWKKGDTFSCWVGHEAL	PLAFTQKTTIDRLAGKPTHVNVSV	489
Db	444	ILRVAEDWKKGDTFSCWVDHEAL	PLAFTQKTTIDRWAGKPTHINVS	500

Search completed: October 25, 2005, 11:07:52  
Job time : 73.272 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 10:59:42 ; Search time 16.842 Seconds  
(without alignments)  
2793.604 Million cell updates/sec

Title: US-10-644-256-3

Perfect score: 2811  
Sequence: 1 MACPGFLWALVISTCLEFSM.....GKPTHNVNVVMAEVDGTCY 489

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1889	72.3	353	1 AIHU	Ig alpha-1 chain C
2	1838	70.4	352	2 S05500	Ig alpha-1 chain C
3	1710.5	65.5	340	2 B2360	Ig alpha-2 chain C
4	1693.5	64.9	340	2 I56230	Ig alpha-2 chain -
5	1656.5	63.4	340	1 A2HU	Ig alpha-2 chain C
6	1275.5	48.9	342	2 I47175	Ig alpha chain C
7	1113	42.6	220	2 C2360	Ig alpha-2 chain C
8	1112.5	42.6	342	2 A45966	Ig alpha chain C
9	1066.5	40.8	344	1 AHMS	Ig alpha chain C
10	1035.5	39.7	357	2 S09267	Ig alpha chain C
11	1022	39.1	357	2 S09269	Ig alpha chain C
12	1016	38.9	357	2 S09265	Ig alpha chain C
13	1002.5	38.4	627	2 S14683	Ig mu chain precu
14	986	37.8	358	2 S09268	Ig alpha chain C
15	984	37.7	347	2 S09274	Ig alpha chain C
16	980	37.5	348	2 S09270	Ig alpha chain C
17	975	37.3	348	2 S09273	Ig alpha chain C
18	970.5	37.2	339	2 S09264	Ig alpha chain C
19	959.5	36.7	343	2 S09272	Ig alpha chain C
20	950	36.4	352	2 S09266	Ig alpha chain C
21	949	36.3	338	2 S09276	Ig alpha chain C
22	943	36.1	360	2 S09271	Ig alpha chain C
23	883.5	33.8	335	2 S09275	Ig alpha chain C
24	842	32.2	585	2 A46507	Ig alpha chain - c
25	838	32.1	299	1 AHRB	Ig mu chain - shee
26	754	28.9	592	2 S25705	Ig gamma-2a chain
27	748	28.6	469	2 S37483	Ig gamma-2b chain
28	743	28.5	475	2 S01321	Ig gamma-2a chain
29	736.5	28.2	446	2 S40295	Ig gamma-2a chain

#### ALIGNMENTS

##### RESULT 1

AIHU

Ig alpha-1 chain C region - human

C/Species: Homo sapiens (man)

C/Date: 22-May-1991 #sequence, revision 03-Oct-1995 #text change 09-Jul-2004

C/Accession: A22360; A92249; A91662; S38979; B53110; A02171

R/Flanagan, J.G.; Lefranc, M.P.; Rabbitts, T.H.

Cell 36, 681-688, 1984

A/Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1 and

A/Reference number: A94653; MUID:84130179; PMID:6421489

A/Accession: A22360

A/Molecule type: DNA

A/Residues: 1-353 <FLA>

A/Cross-references: UNIPROT:P01876

R/Putnam, F.W.; Liu, Y.S.V.; Low, T.L.K.

J. Biol. Chem. 254, 2863-2874, 1979

A/Title: Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal IgA1 protease

A/Reference number: A92249; MUID:79151016; PMID:107164

A/Contents: myeloma protein Bur; disulfide bonds

A/Accession: A92249

A/Molecule type: protein

A/Residues: 1-16, 'Z', '18', 'B', '20', 'B', '22-34', 'Q', '36-45', 'Z', '47-51', 'B', '53-56', 'ZB', '59-61', 'B', '63',

303, 'B', '305-346', 'Q', '348-353 <PUT>

A/Note: this is the final paper in a series  
A/Note: amidation states of residues 178, 197, 238, 239, 243, 244, 287, and 318 were take  
R/Kratzin, H.; Altevogt, P.; Ruban, E.; Kortt, A.; Starosick, K.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 356, 1337-1342, 1975

A/Title: The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.), II: the am

A/Reference number: A91662; MUID:76023781; PMID:809331

A/Accession: A91662

A/Molecule type: protein

A/Residues: 1-16, 'Z', '18', 'B', '20', 'B', '22-34', 'Q', '36-45', 'Z', '47-51', 'B', '53-56', 'ZB', '59-61', 'B', '63',  
'R', '232-237', 'QQ', '240-243', 'Q', '245-283', 'Q', '285-289', 'E', '291-303', 'B', '305-353 <KRA>

R/Experimental source: myeloma protein Tro

R/Fallgreen-Gebauer, E.; Gebauer, W.; Bastian, A.; Kratzin, H.D.; Eiffert, H.; Zimmermann

Biol. Chem. Hoppe-Seyler 374, 1023-1028, 1993

A/Title: The covalent linkage of secretory component to IgA. Structure of sIgA.

A/Reference number: S38978; MUID:94121784; PMID:8292260

A/Accession: S38979

A/Molecule type: protein

A/Residues: 188-196, 'D', '198-201 <FAL>

R/Yang, C.Y.; Kratzin, H.; Gotz, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 360, 1919-1940, 1979

A/Title: Die Primaerstruktur eines monoklonalen IgA1-Immunglobulins (Myelomprotein Tro).

A/Reference number: A91684; MUID:90114124; PMID:393607

A/Contents: annotation; Tro, disulfide bonds

A/Note: Cys-14 bonds to a light chain

R/Calero, M.; Escribano, J.; Grubb, A.; Mendez, E.

J. Biol. Chem. 269, 384-389, 1994

A/Title: Location of a novel type of interpolyptide chain linkage in the human protein

A/Reference number: A53110; MUID:94103241; PMID:7506257

A/Accession: B53110

A;Molecule type: protein  
A;Residues: 346-351,'X',353 <CAL>  
C;Genetics:

A;Gene: GDB:IGHA1

A;Cross-references: GDB:119332; OMIM:146900

A;Map position: 14q32.33-14q32.33

C;Introns: 1/1; 103/1; 223/1

C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds. Immunoglobulin C region; immunoglobulin homology

C;Superfamily: Immunoglobulin C region; immunoglobulin homology

F;140-206/Domain: immunoglobulin homology <IM1>

F;243-315/Domain: immunoglobulin homology <IM2>

F;26-85,77-101,123-180,147-204,250-313/Disulfide bonds: #status experimental

F;105,111,113,119,121/Binding site: carbohydrate (Ser) (covalent) #status experimental

F;122,182/Disulfide bonds: interchain #status experimental

F;144,340/Binding site: carbohydrate (Asn) (covalent) #status experimental

F;192/Disulfide bonds: interchain (to secretory component) (partial) #status experimental

F;192/Binding site: cysteine (Cys) (covalent) (partial) #status experimental

F;192/Disulfide bonds: interchain (partial) #status experimental

F;352/Cross-link: alpha-1-microglobulin-Ig alpha complex chromophore (Cys) (interchain)

F;352/Disulfide bonds: interchain (to J chain) (partial) #status experimental

F;352/Disulfide bonds: interchain (partial) #status experimental

Query Match 72.3%; Score 1889; DB 1; Length 353;

Best Local Similarity 99.4%; Pred. No. 4.7e-94;

Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 137 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVARNPPPSQDAS 196

DB 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVARNPPPSQDAS 60

QY 197 GDIYTTSSQLTLPATOCCLAGKSVTCHVKHYNPSQDVTPCPVPSPPTPTSP 256

DB 61 GDIYTTSSQLTLPATOCCLAGKSVTCHVKHYNPSQDVTPCPVPSPPTPTSP 120

QY 257 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFWTPTSSGKSAVQGPDRDLG 316

DB 121 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFWTPTSSGKSAVQGPDRDLG 180

QY 317 GCYSVSVSLVSGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 376

DB 181 GCYSVSVSLVPGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 240

QY 377 ALNELVTLTCLARGFSPKQVLRWLGQSQELPREKYLTVASROEPSQGTTFPAVTSILRV 436

DB 241 ALNELVTLTCLARGFSPKQVLRWLGQSQELPREKYLTVASROEPSQGTTFPAVTSILRV 300

QY 437 AEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 489

DB 301 AEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353

RESULT 2

S05500

Ig alpha-1 chain C region - Gorilla (fragment)

C;Species: Gorilla gorilla (gorilla)

C;Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 23-Jul-1999

C;Accession: S05500

R;Kawamura, S.; Omoto, K.; Ueda, S.

Nucleic Acids Res. 17, 6732, 1989

A;Title: Nucleotide sequence of the gorilla immunoglobulin alpha 1 gene.

A;Reference number: S05500; MUID:89386006; PMID:2506527

A;Accession: S05500

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-352 <KAN>

A;Cross-references: EMBL:X15045; NID:g22900; PID:CAA33147.1; PID:g22901

A;Introns: 102/1; 222/1

C;Superfamily: immunoglobulin C region; immunoglobulin homology

C;Keywords: Immunoglobulin

F;242-314/Domain: immunoglobulin homology <IM>

Query Match 70.4%; Score 1838; DB 2; Length 352;

Best Local Similarity 96.9%; Pred. No. 2.4e-91;

Matches 341; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 138 SPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVARNPPPSQDAS 197

DB 1 SPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVARNPPPSQDAS 60

QY 198 DLYTSSQLTLPATOCCLAGKSVTCHVKHYNPSQDVTPCPVPSPPTPTSP 257

DB 61 DLYTSSQLTLPATOCCLAGKSVTCHVKHYNPSQDVTPCPVPSPPTPTSP 120

QY 258 CCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFWTPTSSGKSAVQGPDRDLG 317

DB 121 CCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFWTPTSSGKSAVQGPDRDLG 180

QY 318 CYSVSVSLVSGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 377

DB 181 CYSVSVSLVPGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 240

QY 378 LNELVTLTCLARGFSPKQVLRWLGQSQELPREKYLTVASROEPSQGTTFPAVTSILRV 437

DB 241 LNELVTLTCLARGFSPKQVLRWLGQSQELPREKYLTVASROEPSQGTTFPAVTSILRV 300

QY 438 AEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 489

DB 301 AEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 352

RESULT 3

B22360

Ig alpha-2 chain C region (allotype A2m(1)) - human

C;Species: Homo sapiens (man)

C;Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004

C;Accession: B22360

R;Flanagan, J.G.; Lefranc, M.P.; Rabbitts, T.H.

Cell 36, 681-688, 1984

A;Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1 and

A;Reference number: A94653; MUID:84130179; PMID:6421489

A;Accession: B22360

A;Molecule type: DNA

A;Residues: 1-340 <FLA>

A;Cross-references: UNIPROT:P01877

C;Genetics:

A;Gene: GDB:IGHA2

A;Cross-references: GDB:119333; OMIM:147000

A;Map position: 14q32.33-14q32.33

A;Introns: 1/1 103/1; 210/1

C;Superfamily: immunoglobulin C region; immunoglobulin homology

C;Keywords: immunoglobulin

F;230-302/Domain: immunoglobulin homology <IM>

Query Match 65.5%; Score 1710.5; DB 2; Length 340;

Best Local Similarity 90.9%; Pred. No. 1.5e-84;

Matches 321; Conservative 7; Mismatches 12; Indels 13; Gaps 1;

QY 137 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVARNPPPSQDAS 196

DB 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVARNPPPSQDAS 60

QY 197 GDIYTTSSQLTLPATOCCLAGKSVTCHVKHYNPSQDVTPCPVPSPPTPTSP 256

DB 61 GDIYTTSSQLTLPATOCCLAGKSVTCHVKHYNPSQDVTPCPVPSPPTPTSP 107

QY 257 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFWTPTSSGKSAVQGPDRDLG 316

DB 108 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFWTPTSSGKSAVQGPDRDLG 167

QY 317 GCYSVSVSLVSGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 376

DB 168 GCYSVSVSLVPGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 227



A:Cross-references: EMBL:U12594; NID:g555826; PIDN:AAA65943.1; PID:g555827  
C:Gene: IgA2alpha  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:232-304/Domain: immunoglobulin homology <IMM>

Query Match 48.9%; Score 1275.5; DB 2; Length 342;  
Best Local Similarity 68.8%; Pred. No. 2.4e-61;  
Matches 243; Conservative 32; Mismatches 65; Indels 13; Gaps 3;

QY 138 SPTSPKVPFLSLCSTOPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDASG 197  
DB 2 SETSPKIFPLTGLSSPAGYVIACLVRDFEFPSEPLVTWSPSREGVIVRNPPAQ--AG 59  
QY 198 DLYTSSQLTLPATQCLACKSVTCHVKHTNPSQDVTVPVCPSTPTPTSPSTPPSPS 257  
DB 60 GLYTSSQLTLPEQCPAQIUKQVOHLKSSQSVNVPCKV-----LPSDPCCQ 109  
QY 258 CCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFPTWPSGKSAVQGPORDLQCG 317  
DB 110 CCKPSLSLQPPALADLLGSSNASLTCTLSGLKKSEGVFTWQPSGGKDAVQASPTFDSG 169  
QY 318 CYSVSVSLGCAEPNWHGKTFTCTAAYPESKPTLTATLSK-SGNTPRPEVHLLPPPSBEL 376  
DB 170 CYSVSVSLPGCADPMWKGTFSTCTAAHSELKSGALTATITPKVNTPRPQVHLLPPPSBEL 229  
QY 377 ALNELVTLTCLARGSPKDVLRWLGSGOELPREKYLTVASRQEPSOGTTTFAVTSILRV 436  
DB 230 ALNELVTLTCLVRGSPKDVLRWLGSGOELPRDKYLVWESLPEPGCAIPTVAVTSILRV 289  
QY 437 AAEDMKKGDTFSCMVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 489  
DB 290 DAEDMKQDFTSCMVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEAGICY 342

RESULT 7  
C22360  
Ig alpha-2 chain C region (allotype A2m(2)) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 20-Jun-2000  
R:Accession: C22360  
R:Planagan, J.G.; Lefranc, M.P.; Rabbitts, T.H.  
Cell 36, 681-686, 1984  
A:Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1 and alpha-2 chains  
A:Reference number: A94653; MUID:84130179; PMID:6421489  
A:Accession: C22360  
A:Molecule type: DNA  
A:Residues: 1-220 <FLA>  
A:Cross-references: GB:AJ012264; NID:g3819787; PIDN:CAA09968.1; PID:g3819788  
C:Comment: This sequence revises entry A2HU (A02172).

A:Gene: GDB:IGHA2  
A:Cross-references: GDB:119333; OMIM:147000  
A:Map position: 14q32.33-14q32.33  
A:Introns: 90/1  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:110-182/Domain: immunoglobulin homology <IMM>

Query Match 42.6%; Score 1113; DB 2; Length 220;  
Best Local Similarity 93.2%; Pred. No. 7.3e-53;  
Matches 205; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 270 LEDLLGSEANLTCTLTGLRDASGVTFPTWPSGKSAVQGPORDLQCGYSSVSLGCA 329  
DB 1 LEDLLGSEANLTCTLTGLRDASGVTFPTWPSGKSAVQGPORDLQCGYSSVSLGCA 60  
QY 330 EPWNHGKTFTCTAAYPESKPTLTATLSKSGNTFRPEVHLLPPPSBELNELVTLTCLAR 389  
DB 61 QPWNHGETTCTAAPELKTPTANITKSGNTFRPEVHLLPPPSBELNELVTLTCLAR 120  
QY 390 GFSKPDVLRWLGSGOELPREKYLTVASRQEPSOGTTTFAVTSILRVAAEDMKKGDTFSC 449

121 GFSKPDVLRWLGSGOELPREKYLTVASRQEPSOGTTTFAVTSILRVAAEDMKKGDTFSC 180  
QY 450 MYGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 489  
DB 181 MYGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEADGTCY 220

RESULT 8  
A45966  
Ig alpha chain C region - shrew mouse  
C:Species: Mus pahari  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 16-Jul-1999  
R:Accession: A45966  
R:Osborne, B.A.; Golde, T.E.; Schwartz, R.L.; Rudikoff, S.  
Genetics 119, 925-931, 1988  
A:Title: Evolution of the IgA heavy chain gene in the genus Mus.  
A:Reference number: A45966; MUID:88313645; PMID:2842228  
A:Accession: A45966  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-342 <OSB>  
A:Cross-references: GB:X13020  
A:Note: the authors translated the codon AAC for residue 46 as Lys, ATG for residue 334 &  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:232-304/Domain: immunoglobulin homology <IMM>

Query Match 42.6%; Score 1112.5; DB 2; Length 342;  
Best Local Similarity 58.7%; Pred. No. 1.2e-52;  
Matches 209; Conservative 57; Mismatches 71; Indels 19; Gaps 5;

QY 138 SPTSPKVPFLSL---CSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPSQD 194  
DB 2 SPRNTIPLTLPALSEP---VIIGCLIHDFPSGTMNVTWKGSGNDITTVNFPAL- 57  
QY 195 ASGDLYTSSQLTLPATQCLACKSVTCHVKHTNPSQDVTVPVCPSTPTPTSPSTPTP 254  
DB 58 ASGGRYTWSSQLTLPAKECPAGESVKCSVQHDSPNPQELNVKC-----SEPLP 106  
QY 255 SPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFPTWPSGKSAVQGPORD 314  
DB 107 PSTICQPSLSLQRPALDILLGSDASLTCTLSGLKSTEGWFTWETTGKDAVQKFPQD 166  
QY 315 LCGCYSVSVSLGCAEPNWHGKTFTCTAAYPESKPTLTATLSK-SGNTFRPEVHLLPPPS 373  
DB 167 SCGCFSSVSVLPFCACERNWSGASFTCTVTHPSEPLTGTIAKVTNTFPQVHLLPPPS 226  
QY 374 EELALNELVTLTCLARGSPKDVLRWLGSGOELPREKYLTVASRQEPSOGTTTFAVTSI 433  
DB 227 EELALNELVTLTCLVRFPKPELVRLHGHNEELSPESYLVFEPLKEPGEGATTYLVTSV 286  
QY 434 LRVAEDMKKGDTFSCMVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 489  
DB 287 LRVAETWKQDQYSCMVGHEALPMNFTQKTIDRLSGKPTNVSVSVIMSEGDICY 342

RESULT 9  
A45966  
Ig alpha chain C region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Dec-1981 #sequence\_revision 18-Dec-1981 #text\_change 09-Jul-2004  
R:Accession: A91479; A92245; A93857; A02173  
R:Auffray, C.; Nagotte, R.; Sikorav, J.L.; Heidmann, O.; Rougeon, F.  
Gene 13, 365-374, 1981  
A:Title: Mouse immunoglobulin A: nucleotide sequence of the structural gene for the alpha  
A:Reference number: A91479; MUID:81261947; PMID:6790349  
A:Contents: myelomas ABE48 and J558  
A:Accession: A91479  
A:Molecule type: mRNA  
A:Residues: 1-344 <AUF>  
A:Cross-references: UNIPROT:P01878  
R:Robinson, E.A.; Appella, E.  
J. Biol. Chem. 254, 11418-11430, 1979

C;Keywords: immunoglobulin  
F;142-208/Domain: immunoglobulin homology <IMM>

Query Match 39.1%; Score 1035.5; DB 2; Length 357;  
Best Local Similarity 57.1%; Pred. No. 1.6e-48;  
Matches 205; Conservative 43; Mismatches 98; Indels 13; Gaps 8;

Qy 140 TSPKVPPLSL--CSTQPDGNVITA-CLVQGFPPQEPPLSVTWSSGGQVTAARNFPPSQDAS 196  
Db 3 TPPIIFPLTCPGCVLKDTSATIVAGCLIRGFFPRGPLGVTWMDNRANLT---FPPVQSAT 59  
Qy 197 GDLYTTSSQLTLPATOCLAGKSVTCHVKVTNPSQDVTVPVCPV---PSTPPTSPSTP-P 252  
Db 60 SSLYTTCVLSLPAEQCPAGNSVACRVEH--NNKRQDLTVPLACNKRPTIIEPTTKPTCCP 118  
Qy 253 TSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFWTTPSSGKSAVQGGPD 312  
Db 119 CPSPSCCKPSLSLQRPDLGDLDDNSASLTCTLRGLLNPEGAVFTWNTNGKEFVQSOAQ 178  
Qy 313 RDLGCGYSVSSVLSCGAEPMNHGKTFCTTAAYPE-SKTPLTATLSK-SGNTFRPEVHLLP 370  
Db 179 RDHCGCYSVSSVLPSCGAEPMNAGTVTCTVTHPEIDSGSLTATISKDTGSLIPQVHLLP 238  
Qy 371 PSEELALNELVTLTCLARGFSPKDVLRWLOGSQBELPREKYLITWASROEPPSQGTTTFAV 430  
Db 239 PSEELALNALVTLTCLVRGFSKDVLRWYTNKGLQVPKDSFLVWKPLPLPGGQEPPTYAV 298  
Qy 431 TSLIRVAEEDWKGDITFCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 489  
Db 299 TSLIRVPAEDWNQNESYTCVVGHEGLAEHTQKTIDELAGKPTHVNVSVVVADEVGVCY 357

RESULT 11  
S09269  
Ig alpha chain C region - rabbit (fragment)  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 16-Jul-1999  
C;Accession: S09269  
R;Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.  
EMBO J. 8, 4041-4047, 1989  
A;Title: The Iga heavy-chain gene family in rabbit: cloning and sequence analysis  
A;Reference number: S09264; MUID:90076124; PMID:2512120  
A;Accession: S09269  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-357 <BUR>  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: immunoglobulin  
F;142-208/Domain: immunoglobulin homology <IMM>

Query Match 39.1%; Score 1022; DB 2; Length 357;  
Best Local Similarity 56.9%; Pred. No. 8.4e-48;  
Matches 203; Conservative 35; Mismatches 109; Indels 10; Gaps 7;

Qy 140 TSPKVPPLSLCSTQPDGNVIVIAVLQGFPPQEPPLSVTWSSGGQVTAARNFPPSQDASGL 199  
Db 4 TRPILPLSPILPGEPVVGCLIRGFFPRGPLGVTWNTSGENLT---FPPVQSATSSL 60  
Qy 200 YTTSSQLTLPATOCLAGKSVTCHVKVTNPSQDVTVPVCPVSTPPTSPSTP-----PTP 254  
Db 61 YTTCSLLRLLAEQCPENSVACHVEHNYDKGQHVTVSPPECPQPTTPGSDTTTTCPCPCP 120  
Qy 255 SPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFWTTPSSGKSAVQGGPD 314  
Db 121 SPSCGEPSSLQRPFLRDLNNSASLTCTLRGLKQPEGAVFTWEPTNGKNKPVOQSVQSY 180  
Qy 315 LCGCYSVSSVLSCGAEPMNHGKTFCTTAAYPESEK-TPLTATLSK-SGNTFRPEVHLLPPP 372  
Db 181 PCGCVSVSSVLPSCGAEPMNAGTEFTCTVTHPEIEGGPLTAKISKDTGAILPPQVHLLPPP 240  
Qy 373 SEELALNELVTLTCLARGFSPKDVLRWLOGSQBELPREKYLITWASROEPPSQGTTTFAVTS 432  
Db 241 SEELALNELVTLTCLVRGFSKDVLRWYTNKGVNVPENSLVWKPLPLPGGQEPPTYAVTS 300

Matches 199; Conservative 41; Mismatches 105; Indels 14; Gaps 7;  
Best Local Similarity 55.4%; Pred. No. 7e-46;

Qy 141 SPKVFL--SLCSTQPDGNVVIA-CLVQGGFFPQEPPLSVTWSESQGVYARNFPFSDASG 197

```
Db      4  SPRLPPLIHPRCALKDTSATVIAGCLIRGFPPLQPLSVSMNAGKNVT---FPPVPSGTS 60
Qy      198 DLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPSPSTP-----P 252
Db      61  GPYTTCSSLSTPEQCPEDDNNVCHVEHNDKGQNLTVLYP-ECQPTTPSPPTTTTCPCP 119
Qy      253 TSPSCCHPRLSLRPALEDLLLGSEANLCTLTGLRDASGVTTWTTPSSGKSAVQGPDP 312
Db      120 CPSPSCGEPSSLQRPDLGDLNLNSASLCTLRGLLDPEGAFTWEPTFGKEFVQLSPK 179
Qy      313 RDLGCGYSVSVLSCAPWNNHGTFTCTAAYPSK-TPLTATLSK-SGNTFRPEVHLLP 370
Db      180 LDHCGCYSVSVLPGCCAANNAGTKFNCTVTHPEIKGVSLTDIISKDTGVVIAQVHLLP 239
Qy      371 PPSBELALNELVTLTCLARGFSPKDLVRLWLGQSQELPREKYLTVASRQEPSQGTTFPAV 430
Db      240 PPSDELALNALVTLTCLVRGFSKDLVLYWTKGVVPEKDSFLVWKPLPEPGQPTTVAV 299
Qy      431 TSIIRVAAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPTHVNVSVVAEVDGTCY 489
Db      300 TSLLRVPAEDWNQNESYTCVVGHEGLAEHFTQRTIDRLAGRPTHVNVSVVADVEGVCY 358
```

## RESULT 15

```
S09274
Ig alpha chain C region - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09274
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The IgA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference number: S09264; MUID:90076124; PMID:2512120
A:Accession: S09274
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-347 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:132-198/Domain: immunoglobulin homology <IMW>
```

```
Query Match      37.7%; Score 984; DB 2; Length 347;
Best Local Similarity 54.4%; Pred. No. 8.6e-46;
Matches 193; Conservative 45; Mismatches 103; Indels 14; Gaps 4;

Qy      140 TSPKVFPLSLCSTQPDGN---VVIACLVOGFFQEPLEPLSVTWSESGQGVTAARNFPPSQDAS 196
Db      2  TTPGIYPLSLPLRVSDGNSQTVVVGCLIRGFPLGLFVSNVSNRENVSIYNFFPAPTGT 61
Qy      197 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPSPSTPTPSP 256
Db      62 SGPTYACSELILPDTQCLEYDAAACHEYNSVINESLPVPPDP-----CEQCHCP 112
Qy      257 SCCHPRLSLRPALEDLLLGSEANLCTLTGLRDASGVTTWTTPSSGKSAVQGPDPDRDL 316
Db      113 SCEEPSLSLQRPDLRLDLGLSDASLCTLRGLKDPEGAFTWGTNGNPEFVQSPQRPDPC 172
Qy      317 GCYSVSVLSCAPWNNHGTFTCTAAYPSK-TPLTATLSK-SGNTFRPEVHLLPPPE 374
Db      173 GCYSVSVLPGCCAEPWAGTEFTCTVTHPEIEGSLTATISKDTGSLTTPPLVHLLPPPE 232
Qy      375 ELALNELVTLTCLARGFSPKDLVRLWLGQSQELPREKYLTVASRQEPSQGTTFPAVTSIL 434
Db      233 ELALNALVTLTCLVRGFSKDLVLYWTKGVVPEKDSFLVWKPLPEPGQDPTTVAVTSL 292
Qy      435 RVAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPTHVNVSVVAEVDGTCY 489
Db      293 RVPADWNQNESYSCVWAEGHAEHFTQRTIDRLAGRPTHVNVSVVADVEAVCY 347
```

Search completed: October 25, 2005, 11:08:25  
Job time : 18.842 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 10:53:52 ; Search time 80.1449 Seconds  
(without alignments)  
2359.799 Million cell updates/sec

Title: US-10-644-256-3

Perfect score: 2611

Sequence: 1 MACPGFLWLVSTCLEFSM.....GKPTHVNVSVMAEVDGTCY 489

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2374	90.9	497	8	ADP73848 Human ant
2	2335	89.4	630	7	ADF69023 Chloropla
3	2335	89.4	639	7	ADF68991 Chloropla
4	2302	88.2	496	7	ADM05373 Human pro
5	2281.5	87.4	495	7	ADb64852 Human pro
6	2276	87.2	495	7	ADM05400 Human pro
7	2275	87.1	502	7	ADM05560 Human pro
8	2271.5	87.0	496	8	ADR08452 Human pro
9	2258	86.5	456	8	ADP73852 Human ant
10	2230	85.4	495	7	ADM05592 Human pro
11	2217	84.9	492	8	ADR10249 Human pro
12	2215	84.8	494	4	RAb95696 Human pro
13	2215	84.8	494	5	ADil17262 Human NOV
14	2213.5	84.8	496	8	ADQ65888 Novel hum
15	2201.5	84.3	493	4	RAm93283 Human pol
16	2201.5	84.3	493	8	ADL30731 Human pro
17	2200.5	84.3	497	7	ADb65619 Human pro
18	2193.5	84.0	499	8	ADR10340 Human pro
19	2193	84.0	492	8	ADR10164 Human pro
20	2192	84.0	494	7	ADM05393 Human pro
21	2191	83.9	494	8	ADQ65690 Novel hum
22	2190.5	83.9	576	8	ADP69325 Human lun
23	2190	83.9	508	7	ADb65079 Human pro
24	2188.5	83.8	497	7	ADM05559 Human pro
25	2186	83.7	530	4	ABb90574 Human sec

26	2186	83.7	530	5	ABG65440	Abg65440 Human alb
27	2186	83.7	530	8	ADL78707	Adl78707 Albumin f
28	2184.5	83.7	499	8	ADQ65976	Adq65976 Novel hum
29	2179.5	83.5	491	7	ADM05577	Adm05577 Human pro
30	2174	83.3	500	7	ADM05570	Adm05570 Human pro
31	2172.5	83.2	491	7	ADb65089	Adb65089 Human pro
32	2169.5	83.1	497	8	ADQ66283	Adq66283 Novel hum
33	2169	83.1	494	7	ADM05590	Adm05590 Human pro
34	2160.5	82.7	495	8	ADM05429	Adm05429 Human pro
35	2154.5	82.5	500	8	ADR09662	Adr09662 Human pro
36	2149.5	82.3	497	8	ADQ67021	Adq67021 Novel hum
37	2148	82.3	507	3	AAy96304	Aay96304 Human IGF
38	2132	81.7	497	3	AAy44723	Aay44723 Human Imm
39	2125.5	81.4	491	7	ADM05605	Adm05605 Human pro
40	2120.5	81.2	495	4	AAg62159	Aag62159 Human gen
41	2120.5	81.2	495	5	ABG63555	Abg63555 Human alb
42	2120.5	81.2	495	8	ADL76820	Adl76820 Albumin f
43	2120	81.2	686	5	ADil17259	Adil17259 Human NOV
44	2113.5	80.9	475	5	ADil17258	Adil17258 Human NOV
45	2110	80.8	472	5	ADil17260	Adil17260 Human NOV

ALIGNMENTS

RESULT 1

ADP73848  
ID ADP73848 standard; protein; 497 AA.

XX AC ADP73848;

XX DT 09-SEP-2004 (first entry)

XX DE Human anti-HSV antibody HX8 heavy chain (+ mouse leader), SEQ:2.

XX KW Transgenic plant; immunoglobulin production; recombinant production;

XX KW glycosylation; fucose; glycan; virucide; immunotherapy;

XX KW herpes simplex virus; HSV1; HSV2; human; mouse; leader sequence;

XX KW monoclonal antibody HX8; IgA; heavy chain; antibody.

XX OS Homo sapiens.

XX OS Mus sp.

XX OS Chimeric.

XX FH Key Location/Qualifiers

FT Peptide 1..19

FT /note= "Mouse leader sequence (see also SEQ ID NO:4 and 12)";

FT Protein 20..497

FT /note= "Mature HX8 heavy chain"

FT Modified-site 288

FT /note= "N-glycosylated - corresponds to residue Asn269 of the mature heavy chain. This residue is in the CH2 region"

FT Region 476..497

FT /label= Antibody tailpiece

FT /note= "This region contains an Asn residue which is normally fucosylated (see also SEQ ID NO:8). The tailpiece region is absent in certain embodiments"

FT Modified-site 484

FT /note= "Normally N-glycosylated with fucose-containing glycan"

WO2004050838-A2.

PD 17-JUN-2004.

XX 28-NOV-2003; 2003WO-US037905.

XX 27-NOV-2002; 2002US-0429385P.

XX (DOWC ) DOW CHEM CO.

PA (DOWC ) DOW AGROSCIENCES LLC.

PA (EPIC-) EPICYTE PHARM INC.  
XX Briggs K, Glancy T, Hein MB, Hiatt AC, Karnoup AB, Anderson WHK;  
PI Paredy D, Petolino J, Rubin-Wilson B, Taylor D, Roberts JL;  
XX  
XX WPI; 2004-461111/43.  
DR N-PSDB; ADP73847.  
XX  
XX Novel plant-produced immunoglobulin having glycopeptide or glycan profile  
PT with reduced fucosylation, useful for treating herpes simplex virus  
PT infection.  
XX  
XX Claim 74; SEQ ID NO 2; 212pp; English.  
XX  
XX The invention relates to the production of immunoglobulins in plants,  
CC wherein at least a portion of the glycans attached to the immunoglobulins  
CC lack fucose. The immunoglobulins produced can be of any class (i.e., IgG,  
CC IgM, IgD or IgE) and is especially an anti-herpes simplex virus  
CC (HSV) antibody or an anti-alphabeta3, alphabeta5 dual integrin  
CC antibody. The invention also relates to constructs, plasmids and vectors  
CC for producing the immunoglobulins; transformed plant cells, calli, plant  
CC tissues and whole plants for producing the immunoglobulins; methods for  
CC producing the immunoglobulins, the immunoglobulins thus produced; and the  
CC use of such immunoglobulins. The immunoglobulins of the invention may be  
CC used to treat HSV infection or tumour angiogenesis. The invention  
CC provides the advantages of antibody production in plants, such as large  
CC scale production, reduced costs, and elimination of pathogenic  
CC contaminants such as viruses and prions, with a simplified (i.e., non-  
CC plant-specific) glycosylation profile which reduces the risk that the  
CC immunoglobulin may not be functional in animals. The present sequence  
CC represents the heavy chain (with mouse leader sequence) of the human anti  
CC -HSV1/HSV2 monoclonal IgA antibody HX8.  
XX  
XX Sequence 497 AA;

Query Match 90.9%; Score 2374; DB 8; Length 497;  
Best Local Similarity 93.3%; Pred. No. 2,7e-129;  
Matches 446; Conservative 11; Mismatches 11; Indels 10; Gaps 1;

QY 22 QVQLVSGAEVKKPGSSVRVSKASGGTFSSYAIISWVRQAPGGLEWGMGIIPFGTANY 81  
DB 20 QVQLVSGAEVKKPGSSVRVSKASGGTFSSYAIISWVRQAPGGLEWGMGIIPFGTANY 79  
QY 82 AQFQGRVTITADESTAYMELSLRSDETVVYCARDPF-----LHYWGQGLT 131  
DB 80 AQFQDLTITADVSTAYMQLSGLTYEDTAMYYCARVAYMLEPTVTAGGLDVWGQGLT 139  
QY 132 VTVSTASPTSPKVPFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQVTARNFPP 191  
DB 140 VTVSSASPTSPKVPFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQVTARNFPP 199  
QY 192 SQDASGLYTTSSQLTLPATQCLAGKSVTCHVKHTNPDSQDVTVPVSTPTSPSTP 251  
DB 200 SQDASGLYTTSSQLTLPATQCLAGKSVTCHVKHTNPDSQDVTVPVSTPTSPSTP 259  
QY 252 PTPSPCCPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFWTWSSGKSAVQGGP 311  
DB 260 PTPSPCCPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFWTWSSGKSAVQGGP 319  
QY 312 DRDLCCGYSVSSVLSGCAEPWNGKTFCTTAAVPESKTPLTATLSKSGNTFRREVHLLPP 371  
DB 320 ERDLCCGYSVSSVLSGCAEPWNGKTFCTTAAVPESKTPLTATLSKSGNTFRREVHLLPP 379  
QY 372 PSELALNELVTLTCLARGSPKDLVRLVWLGQSOELPREKYLTVASRQEPSCGTTFAVT 431  
DB 380 PSELALNELVTLTCLARGSPKDLVRLVWLGQSOELPREKYLTVASRQEPSCGTTFAVT 439  
QY 432 SILRVAEDWKKGDTFTSCMVGHEALPLAFTQKTDRLAGKPTHVNVSVNAEVDGTCY 489  
DB 440 SILRVAEDWKKGDTFTSCMVGHEALPLAFTQKTDRLAGKPTHVNVSVNAEVDGTCY 497

ADF69023  
ID ADF69023 standard; protein; 630 AA.  
XX  
XX ADF69023;  
AC  
XX  
XX 12-FEB-2004 (first entry)  
DT  
XX  
XX Chloroplast-codon-optimised HSV8-lsc anti-HSV antibody protein.  
DE  
XX  
XX plastid; chloroplast; antibody chimera; tetanus toxin;  
KW herpes simplex virus; HSV; transgenic; HSV8-lsc; large single chain.  
XX  
XX Unidentified.  
OS  
XX Synthetic.  
OS  
XX WO2003091413-A2.  
PN  
XX  
XX 06-NOV-2003.  
PD  
XX  
XX 23-APR-2003; 2003WO-US012997.  
PP  
XX  
XX 23-APR-2002; 2002US-0375129P.  
PR  
XX  
XX 19-DEC-2002; 2002US-0434957P.  
PR  
XX  
XX (SCRI ) SCRIPPS RES INST.  
PA  
XX  
XX Mayfield SP, Franklin S;  
PI  
XX  
XX WPI; 2003-865583/80.  
DR  
XX  
XX N-PSDB; ADF69022.  
DR  
XX  
XX Producing polypeptides in chloroplasts, useful in producing antibodies,  
PT by introducing a first recombinant nucleic acid molecule comprising a  
PT first polynucleotide encoding at least one polypeptide.  
PT  
XX  
XX Claim 176; SEQ ID NO 48; 174pp; English.  
PS  
XX  
XX The invention relates to a novel method for producing a polypeptide in a  
CC plastid comprising introducing a first recombinant nucleic acid molecule  
CC into the plastid where the first recombinant nucleic acid molecule  
CC comprises a first polynucleotide which encodes at least one polypeptide.  
CC The method of the invention may be useful for producing polypeptides in  
CC chloroplasts, specifically antibodies and antibody chimera that bind  
CC tetanus toxin or a herpes simplex virus (HSV). The polynucleotides and  
CC polypeptides may be useful in producing transgenic plants. The current  
CC sequence is that of the chloroplast-codon-optimised HSV8-lsc anti-HSV  
CC antibody protein of the invention.  
XX  
XX Sequence 630 AA;

Query Match 89.4%; Score 2335; DB 7; Length 630;  
Best Local Similarity 92.4%; Pred. No. 6.3e-127;  
Matches 439; Conservative 11; Mismatches 15; Indels 10; Gaps 1;

QY 25 LVQSGAEVKKPGSSVRVSKASGGTFSSYAIISWVRQAPGGLEWGMGIIPFGTANYAQK 84  
DB 134 LEQSGAEVKKPGSSVRVSKASGGTFSSYAIISWVRQAPGGLEWGMGLPIFGTTNVAQK 193  
QY 85 FQGRVTITADESTAYMELSLRSDETVVYCARDPF-----LHYWGQGLT 134  
DB 194 FQDLTITADVSTAYMQLSGLTYEDTAMYYCARVAYMLEPTVTAGGLDVWGQGLT 253  
QY 135 STASPTSPKVPFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQVTARNFPPSD 194  
DB 254 STASPTSPKVPFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQVTARNFPPSD 313  
QY 195 ASGDLYTTSSQLTLPATQCLAGKSVTCHVKHTNPDSQDVTVPVSTPTSPSTPT 254  
DB 314 ASGDLYTTSSQLTLPATQCLAGKSVTCHVKHTNPDSQDVTVPVSTPTSPSTPT 373  
QY 255 SPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFWTWSSGKSAVQGGPDR 314  
DB 374 SPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFWTWSSGKSAVQGGPDR 433

```

QY 315 LCGCYSSVSLGCAEPNHNKFTTCTAAYPEKSTPLTATLSKSGNTFRPEVHLLPPPE 374
DB 434 LCGCYSSVSLPGCAEPNHNKFTTCTAAYPEKSTPLTATLSKSGNTFRPEVHLLPPPE 493
QY 375 ELALNELVTLTCLARGSPKDVLRWLQGSQELPREKYLTVASRQPSQGTTFVAVTSIL 434
DB 494 ELALNELVTLTCLARGSPKDVLRWLQGSQELPREKYLTVASRQPSQGTTFVAVTSIL 553
QY 435 RVAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNVVMAEVDGTCY 489
DB 554 RVAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNVVMAEVDGTCY 608

RESULT 3
ID ADF68991 standard; protein; 639 AA.
XX
AC ADF68991;
XX
DT 12-FEB-2004 (first entry)
XX
DE Chloroplast-codon-optimised single chain anti-HSV antibody protein.
XX
KW plastid; chloroplast; antibody chimera; tetanus toxin;
KW herpes simplex virus; HSV; transgenic.
XX
OS Synthetic.
OS Unidentified.
XX
XX WO2003091413-A2.
XX
XX 06-NOV-2003.
XX
XX 23-APR-2003; 2003WO-US012997.
XX
XX 23-APR-2002; 2002US-0375129P.
XX
XX 19-DEC-2002; 2002US-0434957P.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Mayfield SP, Franklin S;
XX
XX WPI; 2003-865583/80.
XX
XX N-PSDB; ADF68990.
XX
XX Producing polypeptides in chloroplasts, useful in producing antibodies,
XX by introducing a first recombinant nucleic acid molecule comprising a
XX first polynucleotide encoding at least one polypeptide.
XX
XX Claim 176; SEQ ID NO 16; 174pp; English.
XX
XX The invention relates to a novel method for producing a polypeptide in a
XX plastid comprising introducing a first recombinant nucleic acid molecule
XX into the plastid where the first recombinant nucleic acid molecule
XX comprises a first polynucleotide which encodes at least one polypeptide.
XX The method of the invention may be useful for producing polypeptides in
XX chloroplasts, specifically antibodies and antibody chimera that bind
XX tetanus toxin or a herpes simplex virus (HSV). The polynucleotides and
XX polypeptides may be useful in producing transgenic plants. The current
XX sequence is that of the chloroplast-codon-optimised single chain anti-HSV
XX antibody protein of the invention.
XX
XX Sequence 639 AA;
XX
XX Query Match 89.4%; Score 2335; DB 7; Length 639;
XX Best Local Similarity 92.4%; Pred. No. 6.4e-127;
XX Matches 439; Conservative 11; Mismatches 15; Indels 10; Gaps 1;
XX
XX 25 LVQSGAEVKPGSSVRVSKASGCTFSSYAISWVRQAPGQGLEWVGIIPIEGTANYAQK 84
DB 143 LEQSGAEVKPGSSVRVSKASGCTFSSYAISWVRQAPGQGLEWVGIIPIEGTANYAQK 202

```

```

QY 85 FQGRVITADESTAYMELSSLRSEDTAVVYCARDPF-----LHYWGQGLTVTV 134
DB 203 FQDLRITADVSTAYMQLSLGTYEDTAMYYCARVAYMLEPTVTAGCLDVMGKTTTV 262
QY 135 STASPTSPKVPPLSLGSTOPDGNVVIACLVQGFPOEPPLSVTWSESGQGVTAARNPPPSQD 194
DB 263 SPASPTSPKVPPLSLGSTOPDGNVVIACLVQGFPOEPPLSVTWSESGQGVTAARNPPPSQD 322
QY 195 ASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPCVPSTPTPTSPSTPTPT 254
DB 323 ASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPCVPSTPTPTSPSTPTPT 382
QY 255 SPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSVAQGPDPDR 314
DB 383 SPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSVAQGPDPDR 442
QY 315 LCGCYSSVSLGCAEPNHNKFTTCTAAYPEKSTPLTATLSKSGNTFRPEVHLLPPPE 374
DB 443 LCGCYSSVSLPGCAEPNHNKFTTCTAAYPEKSTPLTATLSKSGNTFRPEVHLLPPPE 502
QY 375 ELALNELVTLTCLARGSPKDVLRWLQGSQELPREKYLTVASRQPSQGTTFVAVTSIL 434
DB 503 ELALNELVTLTCLARGSPKDVLRWLQGSQELPREKYLTVASRQPSQGTTFVAVTSIL 562
QY 435 RVAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNVVMAEVDGTCY 489
DB 563 RVAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNVVMAEVDGTCY 617

RESULT 4
ID ADM05373 standard; protein; 496 AA.
XX
AC ADM05373;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein of the invention SEQ ID NO:4058.
XX
KW human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS Homo sapiens.
XX
XX EP1347046-A1.
XX
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
XX
XX 22-MAR-2002; 2002JP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI: 2003-723558/69.
XX N-PSDB; ADM02930.
XX
XX New polynucleotides and polypeptides are useful in gene therapy, for
XX developing a diagnostic marker or medicines for regulating their
XX expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 4058; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
XX polypeptide. A polynucleotide of the invention may have a use in gene
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX as a primer for synthesizing the polynucleotide or as a probe for
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX useful in gene therapy, for developing a diagnostic marker or medicines
XX for regulating their expression and activity, or as a target of gene

```



```
Db 379 PSEELALNELVTLTCLARGFSPKDVLRWLQSGQELPREKYLTVASROBPQSGTTTFAVT 438
Qy 432 SILRVAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 489
Db 439 SILRVAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 496

RESULT 6
ADM05400
XX ADM05400 standard; protein; 495 AA.
AC ADM05400;
XX
XX
XX 20-MAY-2004 (first entry)
XX
XX Human protein of the invention SEQ ID NO:4085.
XX human; gene therapy; diagnostic marker; pharmaceutical.
XX Homo sapiens.
XX
XX EPI347046-A1.
XX
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
XX
XX 22-MAR-2002; 2002JJP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI: 2003-723558/69.
XX N-PSDB; ADM02957.
XX
XX New polynucleotides and polypeptides are useful in gene therapy, for
XX developing a diagnostic marker or medicines for regulating their
XX expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 4085; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
XX polypeptide. A polynucleotide of the invention may have a use in gene
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX as a primer for synthesizing the polynucleotide or as a probe for
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX useful in gene therapy, for developing a diagnostic marker or medicines
XX for regulating their expression and activity, or as a target of gene
XX therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX are useful as pharmaceutical agents. The present sequence represents a
XX protein sequence of the invention.
XX
XX Sequence 495 AA;

Query Match 87.2%; Score 2276; DB 7; Length 495;
Best Local Similarity 89.7%; Pred. No. 1.3e-123;
Matches 427; Conservative 17; Mismatches 24; Indels 8; Gaps 1;

Qy 22 QVQLVQSGAEVKKPGSSVRVSKASGGTSSYAIISWRQAPGQGLEWMGGIPIFGTANY 81
Db 20 QVQLVQSGAEVKKPGAAVKVSKASGDIFTTHLHWVRQAPGQGPENWGMSPSGDRITY 79
Qy 82 AQKFCQGVTTTADSTSTAYNELSSLSDEPTAVNYCARD-----PFLHYWGQGLT 133
Db 80 SQKFRGRVSMTRDSTGTGVNYMELTSLDDPTAVNYCAGEADSAVRTWAFDYWGQGSVT 139
Qy 134 VSTASPTSPKVFPLSLCSTQPDGNWVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFP 193
Db 140 VSSASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFP 199
```

```
Qy 194 DASGDLVTTSSQLTLPATQCLAGKSVTCHVHYTNPSQDVTVPQVPSTPTPTPSPSTPTPT 253
Db 200 DASGDLVTTSSQLTLPATQCLAGKSVTCHVHYTNPSQDVTVPQVPSTPTPTPSPSTPTPT 259
Qy 254 PPSCCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPTSSGKSAVQGPDPDR 313
Db 260 PPSCCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPTSSGKSAVQGPDPDR 319
Qy 314 DLCCGYSSVSVLSCAEPWNHKGTTFTCTAAYPESKTLTATLSKSGNTFRPEVHLLPPPS 373
Db 320 DLCCGYSSVSVLPCCAEPWNHKGTTFTCTAAYPESKTLTATLSKSGNTFRPEVHLLPPPS 379
Qy 374 EELALNELVTLTCLARGFSPKDVLRWLQSGQELPREKYLTVASROBPQSGTTTFAVTSI 433
Db 380 EELALNELVTLTCLARGFSPKDVLRWLQSGQELPREKYLTVASROBPQSGTTTFAVTSI 439
Qy 434 LRVAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 489
Db 440 LRVAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 495

RESULT 7
ADM05560
XX ADM05560 standard; protein; 502 AA.
XX
XX AC ADM05560;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human protein of the invention SEQ ID NO:4245.
XX
XX human; gene therapy; diagnostic marker; pharmaceutical.
XX Homo sapiens.
XX
XX EPI347046-A1.
XX
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
XX
XX 22-MAR-2002; 2002JJP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI: 2003-723558/69.
XX N-PSDB; ADM03117.
XX
XX New polynucleotides and polypeptides are useful in gene therapy, for
XX developing a diagnostic marker or medicines for regulating their
XX expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 4245; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
XX polypeptide. A polynucleotide of the invention may have a use in gene
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX as a primer for synthesizing the polynucleotide or as a probe for
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX useful in gene therapy, for developing a diagnostic marker or medicines
XX for regulating their expression and activity, or as a target of gene
XX therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX are useful as pharmaceutical agents. The present sequence represents a
XX protein sequence of the invention.
XX
XX Sequence 502 AA;

Query Match 87.1%; Score 2275; DB 7; Length 502;
Best Local Similarity 87.2%; Pred. No. 1.5e-123;
```

Matches 435; Conservative 16; Mismatches 30; Indels 18; Gaps 3;	
QY	6 FLWALVISTCLFMSAQVLVQSGAEVKKPGSSVRVSCSKAGGTFFSSYALSWVRQAPQG 65
Db	7 FLFVAAATGVR---AQVHMQSAAAEVKKPGSSVRVSCRTPTSPNNHAIWSVRQAPGRG 63
QY	66 LEWGGIIPFGTANYAOKFGQRTVITADESTAYMELSSLRSEDTAVYCA-RDP--- 121
Db	64 PEMGDRVPELTKTTTAOKFLGRVITADESTAYMDLTSLSYEDTAMVYACGRDEEYC 123
QY	122 -----FLHYMGQGTTLVTVSTASPTSPKVFPPLSLCSTQPDGNNVVIACLVQGFPPQ 170
Db	124 IDLSCTSGVYGLNGQGTMMVSSASPTSPKVFPPLSLCSTQPDGNNVVIACLVQGFPPQ 163
QY	171 EPLSVTWSESGQGVTVARNPPPSQDAGDLYTTSSTQLTPATQCLAGKSVTCHVKHYTNPS 230
Db	184 EPLSVTWSESGQGVTVARNPPPSQDAGDLYTTSSTQLTPATQCLAGKSVTCHVKHYTNPS 243
QY	231 QDVTVPVCPVPTPTPTSPPTSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRD 290
Db	244 QDVTVPVCPVPTPTPTSPPTSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRD 303
QY	291 ASGVTTFTWTPSSGKSAVQGPDRDLCCGYSVSVSLSGCAEPNMHNGKTFCTAAYPESKTP 350
Db	304 ASGVTTFTWTPSSGKSAVQGPDRDLCCGYSVSVSLPGCAEPNMHNGKTFCTAAYPESKTP 363
QY	351 LTATLSKSGNTFRPEVHLLPPPESEALNELVLTCLARGFSPKDVLRVWLQSQBELPRE 410
Db	364 LTATLSKSGNTFRPEVHLLPPPESEALNELVLTCLARGFSPKDVLRVWLQSQBELPRE 423
QY	411 KYLTWASROEPSGTTTFAVTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAG 470
Db	424 KYLTWASROEPSGTTTFAVTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAG 483
QY	471 KPTHNVSVVMAEVDGTCY 489
Db	484 KPTHNVSVVMAEVDGTCY 502
RESULT 8	
ADP08452	
ID	ADP08452 standard; protein; 496 AA.
XX	
AC	ADP08452;
XX	
DT	04-NOV-2004 (first entry)
DE	Human protein useful for treating neurological disease Seq 1958.
XX	
KW	human; oligo-capping method; diagnostic marker; gene therapy;
KW	osteoporosis; neurological disease; Alzheimer's disease;
KW	Parkinson's disease; dementia; short memory; cancer;
KW	sense or motor function; emotional reaction; fear response; panic;
KW	osteopathic; neuroprotective; nootropic; antiparkinsonian; cytosstatic;
KW	tranquilliser.
XX	
OS	Homo sapiens.
XX	
PN	EP1447413-A2.
XX	
PD	18-AUG-2004.
XX	
XX	12-FEB-2004; 2004EP-00003145.
PF	
XX	
PR	14-FEB-2003; 2003JP-00102207.
PR	09-MAY-2003; 2003JP-00131452.
XX	
XX	(REAS-) RES ASSOC BIOTECHNOLOGY.
XX	
PI	Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI	Wakamatsu A, Ishii S, Nagai K, Irie R;
XX	
PR	WPI; 2004-583265/57.

DR	N-PSDB; ADR06496.
XX	
PT	New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT	Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX	
PS	Claim 1; SEQ ID NO 1958; 2686pp; English.
XX	
CC	This invention relates to novel, isolated full length human cDNA
CC	molecules and the encoded proteins thereof. Specifically, it refers to
CC	cDNA clones obtained by an oligo-capping method, where none of these
CC	clones are identical to any known human mRNAs. The present invention
CC	describes an immunoassay to identify agonists and antagonists, as well as
CC	antibodies, antisense molecules and siRNAs that can all be used to bind
CC	to and modulate expression of the cDNA molecules. As such, these
CC	molecules are useful for diagnostic markers or therapeutic targets for
CC	the various diseases or morbid states. In particular, they are useful in
CC	gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC	disease, Parkinson's disease, dementia, short memory and various cancers,
CC	as well as for maintaining equilibrium of sense or motor function, and
CC	for treating emotional reaction, fear response and panic. Accordingly,
CC	they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC	cytostatic and tranquilliser activities. This polypeptide is a protein
CC	encoded by a full length human cDNA sequence of the invention. NOTE: This
CC	sequence is not given in the sequence listing of the specification but
CC	can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC	office.
XX	
QQ	Sequence 496 AA;
XX	
Query Match 87.0%; Score 2271.5; DB 8; Length 496;	
Best Local Similarity 89.1%; Pred. No. 2,3e-123;	
Matches 426; Conservative 14; Mismatches 29; Indels 9; Gaps 1;	
QY	21 AQVLVQSGAEVKKPGSSVRVSCSKAGGTFFSSYALSWVRQAPQGLEWGGIIPFGTAN 80
Db	19 SQAQLVQSGAEAKKPGASVKISCKAGYPSGVLHWRQAQPGQGLEWLGISAGYDATK 78
QY	81 YAQKFGQRTVITADESTAYMELSSLRSEDTAVYCARDPPFLHY-----WGQGT 131
Db	79 YSQRFOQRTITRDTASATVYLEWSSLTSDDTAVLYCAREGDEDEDYDGLGAFDWMQGT 138
QY	132 VTVSTASPTSPKVFPPLSLCSTQPDGNNVVIACLVQGFPOEPLSVTWSESGQVTARNFPP 191
Db	139 VTVSPASPTSPKVFPPLSLCSTQPDGNNVVIACLVQGFPOEPLSVTWSESGQVTARNFPP 198
QY	192 SQDASGDLYTTSSTQLTPATQCLAGKSVTCHVKHYTNPSQDVTVPVCPVPTPTSPSTP 251
Db	199 SQDASGDLYTTSSTQLTPATQCLAGKSVTCHVKHYTNPSQDVTVPVCPVPTPTSPSTP 258
QY	252 PTPSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTTWTPSSGKSAVOGPP 311
Db	259 PTPSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTTWTPSSGKSAVOGPP 318
QY	312 DRDLCCGYSVSVSLSGCAEPNMHNGKTFCTAAYPESKTPLTATLSKSGNTFRREVHLLPP 371
Db	319 DRDLCCGYSVSVSLPGCAEPNMHNGKTFCTAAYPESKTPLTATLSKSGNTFRREVHLLPP 378
QY	372 PSEELALNELVLTCLARGFSPKDVLRVWLQSQBELPREKYLWASRQPSQGTTFITFAVT 431
Db	379 PSEELALNELVLTCLARGFSPKDVLRVWLQSQBELPREKYLWASRQPSQGTTFITFAVT 438
QY	432 SILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHNVSVVMAEVDGTCY 489
Db	439 SILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHNVSVVMAEVDGTCY 496
RESULT 9	
ADP73852	
ID	ADP73852 standard; protein; 456 AA.
XX	
AC	ADP73852;
XX	
DT	09-SEP-2004 (first entry)

XX DE Human anti-HSV antibody HX8 heavy chain (without tailpiece), SEQ:6.  
 XX KW Transgenic plant; immunoglobulin production; recombinant production;  
 KW glycosylation; fucose; glycan; virucide; immunotherapy;  
 KW herpes simplex virus; HSV1; HSV2; human; monoclonal antibody HX8; IgA;  
 KW heavy chain; antibody.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 288  
 FT /note= "N-glycosylated. This residue is in the CH2  
 FT region"  
 XX WO2004050838-A2.  
 XX PD 17-JUN-2004.  
 XX PF 28-NOV-2003; 2003WO-US037905.  
 XX PR 27-NOV-2002; 2002US-0429385P.  
 XX PA (DOWC ) DOW CHEM CO.  
 PA (DOWC ) DOW AGROSCIENCES LLC.  
 PA (EPIC-) EPICYTE PHARM INC.  
 XX PI Briggs K, Glancy T, Hein MB, Hiatt AC, Karnoup AL, Anderson WHK;  
 PI Pareddy D, Petolino J, Rubin-Wilson B, Taylor D, Roberts JL;  
 XX WPI; 2004-461111/43.  
 DR N-PSDB; ADP73851.  
 XX Novel plant-produced immunoglobulin having glycopeptide or glycan profile  
 PT with reduced fucosylation, useful for treating herpes simplex virus  
 PT infection.  
 XX Claim 75; SEQ ID NO 6; 212pp; English.  
 XX The invention relates to the production of immunoglobulins in plants,  
 CC wherein at least a portion of the glycans attached to the immunoglobulins  
 CC lack fucose. The immunoglobulins produced can be of any class (i.e., IgG,  
 CC IgA, IgM, IgE or IgD) and is especially an anti-herpes simplex virus  
 CC (HSV) antibody or an anti-alpha/beta3, alpha/beta5 dual integrin  
 CC antibody. The invention also relates to constructs, plasmids and vectors  
 CC for producing the immunoglobulins; transformed plant cells, calli, plant  
 CC tissues and whole plants for producing the immunoglobulins; methods for  
 CC producing the immunoglobulins, the immunoglobulins thus produced; and the  
 CC use of such immunoglobulins. The immunoglobulins of the invention may be  
 CC used to treat HSV infection or tumour angiogenesis. The invention  
 CC provides the advantages of antibody production in plants, such as large  
 CC scale production, reduced costs, and elimination of pathogenic  
 CC contaminants such as viruses and prions, with a simplified (i.e., non-  
 CC plant-specific) glycosylation profile which reduces the risk that the  
 CC immunoglobulin may not be functional in animals. The present sequence  
 CC represents the heavy chain (without tailpiece) of the human anti-  
 CC HSV1/HSV2 monoclonal IgA antibody HX8.  
 XX Sequence 456 AA;  
 SQ

Query Match  
 Best Local Similarity 86.5%; Score 2258; DB 8; Length 456;  
 Matches 424; Conservative 11; Mismatches 11; Indels 10; Gaps 1;  
 QY 22 QVQLVQSGAEVKKPGSSVRVSCKASGGTFSSYAISSWVROAPGQGLEWMGGIPIFGTANY 81  
 DB 1 QVQLVQSGAEVKKPGSSVRVSCKASGGSSFSYAINWVRQAPGQGLEWMGLMPFGTNY 60  
 QY 82 AQKFGQRTITADESTAYNELSSRLSRDEDAVYYCARDPF-----LHYWGQGTLL 131  
 DB 61 AQKQDRLLTITADVSTSTAYVQLSGLTYEDTAMYYCARVAMLSPTVTAGGLDWGQGTLL 120  
 QY 132 VTVSTASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESQGVTVARNFPP 191

Db 121 VTVSSASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESQGVTVARNFPP 180  
 QY 192 SODASGDLTYTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPCVPVPTPTPSPSTP 251  
 Db 181 SODASGDLTYTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPCVPVPTPTPSPSTP 240  
 QY 252 PTPSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTTPTPSSGKSAVQGGP 311  
 Db 241 PTPSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTTPTPSSGKSAVQGGP 300  
 QY 312 DRDLCCGYSSVSVLSCAEPNHNKTFCTTAAYPESKTLTATLSKSGNTRPRPEVHLLPP 371  
 Db 301 ERDLCCGYSSVSVLPGCAEPNHNKTFCTTAAYPESKTLTATLSKSGNTRPRPEVHLLPP 360  
 QY 372 PSEELALNELVTLTCLARGFSPKDVLRWLGOSQBELPREKYLTVASRQEPSQGTITFAVT 431  
 Db 361 PSEELALNELVTLTCLARGFSPKDVLRWLGOSQBELPREKYLTVASRQEPSQGTITFAVT 420  
 QY 432 SILRVAEDWKKGDTFSCMGHEALPLAFTQKTIDR 467  
 Db 421 SILRVAEDWKKGDTFSCMGHEALPLAFTQKTIDR 456  
 RESULT 10  
 ADM05592  
 ID ADM05592 standard; protein; 495 AA.  
 XX AC ADM05592;  
 XX DT 20-MAY-2004 (first entry)  
 XX Human protein of the invention SEQ ID NO:4277.  
 DE human; gene therapy; diagnostic marker; pharmaceutical.  
 KW Homo sapiens.  
 OS EPI347046-A1.  
 FN 24-SEP-2003.  
 PD 12-APR-2002; 2002EP-00008400.  
 PF 22-MAR-2002; 2002JP-00137785.  
 PR (REAS-) RES ASSOC BIOTECHNOLOGY.  
 PA Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX WPI; 2003-723558/69.  
 DR N-PSDB; ADM03149.  
 XX New polynucleotides and polypeptides are useful in gene therapy, for  
 PT developing a diagnostic marker or medicines for regulating their  
 PT expression and activity, or as a target of gene therapy.  
 XX Claim 1; SEQ ID NO 4277; 305pp; English.  
 XX The invention relates to a novel human polynucleotide and the encoded  
 CC polypeptide. A polynucleotide of the invention may have a use in gene  
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
 CC as a primer for synthesizing the polynucleotide or as a probe for  
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03759 are  
 CC useful in gene therapy, for developing a diagnostic marker or medicines  
 CC for regulating their expression and activity, or as a target of gene  
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
 CC are useful as pharmaceutical agents. The present sequence represents a  
 CC protein sequence of the invention.  
 XX Sequence 495 AA;  
 SQ





PR 15-MAR-2001; 2001US-0276448P.  
 PR 15-MAR-2001; 2001US-0276450P.  
 PR 16-MAR-2001; 2001US-0276397P.  
 PR 16-MAR-2001; 2001US-0276768P.  
 PR 20-MAR-2001; 2001US-0278652P.  
 PR 26-MAR-2001; 2001US-0278775P.  
 PR 26-MAR-2001; 2001US-0278778P.  
 PR 29-MAR-2001; 2001US-0279882P.  
 PR 29-MAR-2001; 2001US-0279884P.  
 PR 30-MAR-2001; 2001US-0280147P.  
 PR 11-APR-2001; 2001US-0282992P.  
 PR 11-APR-2001; 2001US-0283083P.  
 PR 20-APR-2001; 2001US-0285133P.  
 PR 23-APR-2001; 2001US-0285749P.  
 PR 03-MAY-2001; 2001US-0288327P.  
 PR 03-MAY-2001; 2001US-0288504P.  
 PR 29-MAY-2001; 2001US-0294047P.  
 PR 30-MAY-2001; 2001US-0294473P.  
 PR 08-JUN-2001; 2001US-0296964P.  
 PR 18-JUN-2001; 2001US-0298959P.  
 PR 19-JUN-2001; 2001US-0299324P.  
 PR 13-AUG-2001; 2001US-0312020P.  
 PR 16-AUG-2001; 2001US-0312889P.  
 PR 16-AUG-2001; 2001US-0312908P.  
 PR 21-AUG-2001; 2001US-0313390P.  
 PR 28-AUG-2001; 2001US-0315470P.  
 PR 31-AUG-2001; 2001US-0316447P.  
 PR 07-SEP-2001; 2001US-0318115P.  
 PR 07-SEP-2001; 2001US-0318118P.  
 PR 12-SEP-2001; 2001US-0318740P.  
 PR 19-SEP-2001; 2001US-0323379P.  
 PR 18-OCT-2001; 2001US-0330245P.  
 PR 18-OCT-2001; 2001US-0330308P.  
 PR 14-NOV-2001; 2001US-0332701P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 XX Li L, Gangolli EA, Padigaru M, Anderson DM, Rastelli L, Miller CE;  
 PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;  
 PI Furtak K, Grose WM, Alsobrook JP, Lepley DM, Rieger DK, Burgees CB;  
 XX WPI; 2002-706998/76.  
 XX  
 XX New NOVX polypeptides and nucleic acids, useful for preventing or  
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 XX Disclosure; SEQ ID NO 798; 1498pp; English.  
 XX  
 CC This invention relates to a novel nucleic acids, and encoded polypeptides  
 CC thereof, which have properties related to the stimulation of biochemical  
 CC or physiological responses in a cell, tissue, organ or organism.  
 CC Specifically, it refers to the use of biologically active fragments for  
 CC diagnostic and prognostic assays and furthermore in the treatment of  
 CC diverse pathological conditions. The present invention describes novel  
 CC human and murine NOVX proteins, as well as methods to modulate their  
 CC expression using antisense oligos, ribozymes and peptide nucleic acids.  
 CC The NOVX polypeptides, polynucleotides and antibodies are useful in  
 CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,  
 CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in  
 CC treating or preventing diseases such as inflammation, autoimmune  
 CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
 CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,  
 CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
 CC and epilepsy. Accordingly, these molecules have many activities including  
 CC cytotatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,  
 CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,  
 CC antiasthmatic, nephrotropic, antiatheritic, hepatotropic,  
 CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
 CC relaxant and anticonvulsant. In addition, they are useful in screening  
 CC assays to identify small molecules that modulate or inhibit, for example,  
 CC

CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
 CC used as in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
 CC of the invention.  
 XX  
 XX Sequence 494 AA;  
 Query Match 84.8%; Score 2215; DB 5; Length 494;  
 Best Local Similarity 84.8%; Pred. No. 4.2e-120;  
 Matches 417; Conservative 26; Mismatches 41; Indels 8; Gaps 2;  
 QY 5 GFLWALVISTCLEFSSMAQVQLVQSGAEVKKPGSSVRVSCKASGCTFSSVAISVVRAPQG 64  
 Db 4 GLRWVFLVA-FLQGVQCEVQLVQSGGLVKPGSLRLSCAASGLSTSTYAMNVRQAPGK 62  
 QY 65 GLEWMGGIIPFGTANYAOKFQGRVTITADESTSTAYMELSSLRSDTAVVYCARDP-- 122  
 Db 63 GLEWVSSISRSDIYYRDSVKGRFTISRDNKNSLYLQVNSLRVDDTAVVYCARDSCNG 122  
 QY 123 -----LHYNGQGTLVTVSTASPTSPKVFPLSLCSTQPDGNNVIACLVQGFPPQEPUSVTW 177  
 Db 123 AICYGSPMGQGTLVTVSSASPTSPKVFPLSLCSTQPDGNNVIACLVQGFPPQEPUSVTW 182  
 QY 178 SESGQGVTAARNFPPSQDASGDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVP 237  
 Db 183 SESGQGVTAARNFPPSQDASGDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVP 242  
 QY 238 PVPSTPTPTSPPTPTPTSPSCCHPRLSLHRPALEDLLGSEANLTCITGLRLDASGVTF 297  
 Db 243 PVPSTPTPTSPPTPTPTSPSCCHPRLSLHRPALEDLLGSEANLTCITGLRLDASGVTF 302  
 QY 298 WTPSSGKSAVQGPDPDRDLGCGYSVSVLSGCAEPWNHKGTFCTCTAAYPESKTPLTATLSK 357  
 Db 303 WTPSSGKSAVQGPDPDRDLGCGYSVSVLSGCAEPWNHKGTFCTCTAAYPESKTPLTATLSK 362  
 QY 358 SGNTRFPEVHLPPPPSEELALNELVLTLCIARGFSPKDVLRVWLOGSQELPREKYLTTAS 417  
 Db 363 SGNTRFPEVHLPPPPSEELALNELVLTLCIARGFSPKDVLRVWLOGSQELPREKYLTTAS 422  
 QY 418 ROEPPSGTTFFAVTSILRVAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNV 477  
 Db 423 ROEPPSGTTFFAVTSILRVAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNV 482  
 QY 478 SVVMAEVDGTCY 489  
 Db 483 SVVMAEVDGTCY 494  
 RESULT 14  
 ADQ65888  
 ID ADQ65888 standard; protein; 496 AA.  
 XX  
 AC ADQ65888;  
 XX  
 DT 07-OCT-2004 (first entry)  
 XX  
 DE Novel human protein sequence #861.  
 XX  
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
 KW gene therapy; diagnostic marker; morbid state; osteoporosis;  
 KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;  
 KW cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1440981-A2.  
 XX  
 PD 28-JUL-2004.  
 XX  
 PP 21-JAN-2004; 2004EP-00001196.  
 XX  
 PP 21-JAN-2003; 2003JP-00102206.  
 PR 09-MAY-2003; 2003JP-00131392.





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:07:58 ; Search time 94.7482 Seconds  
(without alignments)  
1555.504 Million cell updates/sec

Title: US-10-644-256-3\_COPY\_137\_489

Perfect score: 1898  
Sequence: 1 ASPTSPKVFPLSLCSTQPDG.....GRPTHNVSVYMAEVDGTCY 353

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1862994 seqs, 417510619 residues  
Total number of hits satisfying chosen parameters: 1862994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1898	100.0	489	17	US-10-644-256-3
2	1893	99.7	494	15	US-10-108-260A-4078
3	1893	99.7	495	11	US-09-833-245-302
4	1893	99.7	497	15	US-10-104-047-3773
5	1893	99.7	508	15	US-10-108-260A-3028
6	1889	99.5	353	13	US-10-047-542-16
7	1889	99.5	384	15	US-10-072-012-797
8	1889	99.5	393	14	US-10-221-945-3
9	1889	99.5	491	15	US-10-104-047-3243
10	1889	99.5	491	15	US-10-108-260A-4262
11	1889	99.5	491	15	US-10-108-260A-4290

12	1889	99.5	494	15	US-10-108-260A-4275
13	1889	99.5	494	15	US-10-072-012-798
14	1889	99.5	495	15	US-10-108-260A-4085
15	1889	99.5	495	15	US-10-108-260A-4114
16	1889	99.5	496	15	US-10-104-047-3006
17	1889	99.5	496	15	US-10-108-260A-4058
18	1889	99.5	500	15	US-10-108-260A-4684
19	1889	99.5	502	15	US-10-108-260A-4245
20	1889	99.5	508	15	US-10-104-047-3233
21	1889	99.5	630	15	US-10-422-628-48
22	1889	99.5	639	15	US-10-422-628-16
23	1888	99.5	353	20	US-11-003-819-55
24	1885	99.3	495	15	US-10-108-260A-4277
25	1885	99.3	497	15	US-10-108-260A-4244
26	1882	99.2	500	15	US-10-108-260A-4255
27	1878.5	99.0	354	17	US-10-872-932A-32
28	1878.5	99.0	354	18	US-10-810-881A-31
29	1867	98.4	530	9	US-09-800-729-112
30	1867	98.4	530	11	US-09-833-245-2189
31	1862	98.1	494	9	US-09-800-729-216
32	1828	96.3	686	15	US-10-072-012-795
33	1819	95.8	472	15	US-10-072-012-796
34	1811	95.4	475	15	US-10-072-012-794
35	1710.5	90.1	340	13	US-10-047-542-18
36	1710.5	90.1	340	17	US-10-872-932A-33
37	1710.5	90.1	340	18	US-10-810-881A-32
38	1710.5	90.1	487	9	US-09-800-729-145
39	1710.5	90.1	487	11	US-09-833-245-2194
40	1681.5	88.6	538	13	US-10-047-542-99
41	1681.5	88.6	669	9	US-09-807-721-2
42	1676.5	88.3	799	13	US-10-047-542-8
43	1676.5	88.3	822	13	US-10-047-542-48
44	1599	84.2	338	15	US-10-072-012-322
45	1350.5	71.2	520	14	US-10-207-655-286

ALIGNMENTS

RESULT 1  
US-10-644-256-3  
; Sequence 3, Application US/10644256  
; Publication No. US20050106722A1  
; GENERAL INFORMATION:  
; APPLICANT: Jones, David HA  
; APPLICANT: Bout, Abraham  
; TITLE OF INVENTION: Efficient Production of IgA in Recombinant Mammalian Cells  
; FILE REFERENCE: 2578-6077  
; CURRENT APPLICATION NUMBER: US/10/644,256  
; CURRENT FILING DATE: 2003-08-20  
; PRIOR APPLICATION NUMBER: US 09/549,463  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: US 60/129,452  
; PRIOR FILING DATE: 1999-04-15  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: artificial  
; FEATURE:  
; OTHER INFORMATION: Amino acid sequence anti-EpCAM IgA heavy chain  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(21)  
; OTHER INFORMATION: leader peptide  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (22)..(136)  
; OTHER INFORMATION: VH Region  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (137)..(238)

```

; OTHER INFORMATION: CH1 Region
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (239)..(359)
; OTHER INFORMATION: CH2 Region
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (360)..(489)
; OTHER INFORMATION: CH3 Region
US-10-644-256-3

Query Match      100.0%; Score 1898; DB 17; Length 489;
Best Local Similarity 100.0%; Pred. No. 3.6e-117;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGOGVTARNPPPSQDAS 60
Db 137 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGOGVTARNPPPSQDAS 196

QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVCPVSTPTPTSPSTPTPSP 120
Db 197 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVCPVSTPTPTSPSTPTPSP 256

QY 121 SCCHPRLSHRPALEDLLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 180
Db 257 SCCHPRLSHRPALEDLLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 316

QY 181 GCYSVSSVLSGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTRPRPEVHLLPPPSEEL 240
Db 317 GCYSVSSVLSGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTRPRPEVHLLPPPSEEL 376

QY 241 ALNELVTLTCLARGFSPKDVLRWLGSOELPREKYLTVASROEPSQGTTFPAVTSILRV 300
Db 377 ALNELVTLTCLARGFSPKDVLRWLGSOELPREKYLTVASROEPSQGTTFPAVTSILRV 436

QY 301 AAEDWKKGDTFSCMGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 353
Db 437 AAEDWKKGDTFSCMGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 489

RESULT 2
US-10-108-260A-4078
; Sequence 4078, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4078
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4078

Query Match      99.7%; Score 1893; DB 15; Length 494;
Best Local Similarity 99.7%; Pred. No. 7.7e-117;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGOGVTARNPPPSQDAS 60
Db 142 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGOGVTARNPPPSQDAS 201

QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVCPVSTPTPTSPSTPTPSP 120
Db 202 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVCPVSTPTPTSPSTPTPSP 261

QY 121 SCCHPRLSHRPALEDLLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 180
Db 262 SCCHPRLSHRPALEDLLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 321
```

```

QY 181 GCYSVSSVLSGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTRPRPEVHLLPPPSEEL 240
Db 322 GCYSVSSVLPGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTRPRPEVHLLPPPSEEL 381

QY 241 ALNELVTLTCLARGFSPKDVLRWLGSOELPREKYLTVASROEPSQGTTFPAVTSILRV 300
Db 382 ALNELVTLTCLARGFSPKDVLRWLGSOELPREKYLTVASROEPSQGTTFPAVTSILRV 441

QY 301 AAEDWKKGDTFSCMGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 353
Db 442 AAEDWKKGDTFSCMGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 494

RESULT 3
US-09-833-245-302
; Sequence 302, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 302
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-302

Query Match      99.7%; Score 1893; DB 11; Length 495;
Best Local Similarity 99.7%; Pred. No. 7.8e-117;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGOGVTARNPPPSQDAS 60
Db 143 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGOGVTARNPPPSQDAS 202

QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVCPVSTPTPTSPSTPTPSP 120
Db 203 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVCPVSTPTPTSPSTPTPSP 262

QY 121 SCCHPRLSHRPALEDLLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 180
Db 263 SCCHPRLSHRPALEDLLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 322

QY 181 GCYSVSSVLSGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTRPRPEVHLLPPPSEEL 240
Db 323 GCYSVSSVLPGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTRPRPEVHLLPPPSEEL 382

QY 241 ALNELVTLTCLARGFSPKDVLRWLGSOELPREKYLTVASROEPSQGTTFPAVTSILRV 300
Db 383 ALNELVTLTCLARGFSPKDVLRWLGSOELPREKYLTVASROEPSQGTTFPAVTSILRV 442

QY 301 AAEDWKKGDTFSCMGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 353
Db 443 AAEDWKKGDTFSCMGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 495

RESULT 4
US-10-104-047-3773
; Sequence 3773, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
```

; TITLE OF INVENTION: No. US20030236392A1el full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3773  
; LENGTH: 497  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-3773

Query Match 99.7%; Score 1893; DB 15; Length 497;  
Best Local Similarity 99.7%; Pred. No. 7.8e-117;  
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60  
Db 145 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 204  
Qy 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYNPSQDVTVPCEVPSPTPTPTPTPTPTPTSP 120  
Db 205 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYNPSQDVTVPCEVPSPTPTPTPTPTPTSP 264  
Qy 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 180  
Db 265 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 324  
Qy 181 GCYSVSVSLGCAEPWNHKGFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Db 325 GCYSVSVSLGCAEPWNHKGFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 384  
Qy 241 ALNELVTLTCLARGFSPKDVLRWLQSQELPREKYLWTASRQPSQGTTFFAVTSILRV 300  
Db 385 ALNELVTLTCLARGFSPKDVLRWLQSQELPREKYLWTASRQPSQGTTFFAVTSILRV 444  
Qy 301 AAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 445 AAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 497

RESULT 5  
US-10-108-260A-3028  
; Sequence 3028, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3028  
; LENGTH: 508  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3028

Query Match 99.7%; Score 1893; DB 15; Length 508;  
Best Local Similarity 99.7%; Pred. No. 8e-117;  
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60  
Db 156 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 215  
Qy 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYNPSQDVTVPCEVPSPTPTPTPTPTPTPTSP 120  
Db 216 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYNPSQDVTVPCEVPSPTPTPTPTPTPTPTSP 275

Qy 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 180  
Db 276 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 335  
Qy 181 GCYSVSVSLGCAEPWNHKGFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Db 336 GCYSVSVSLGCAEPWNHKGFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 395  
Qy 241 ALNELVTLTCLARGFSPKDVLRWLQSQELPREKYLWTASRQPSQGTTFFAVTSILRV 300  
Db 396 ALNELVTLTCLARGFSPKDVLRWLQSQELPREKYLWTASRQPSQGTTFFAVTSILRV 455  
Qy 301 AAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 456 AAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 508

RESULT 6  
US-10-047-542-16  
; Sequence 16, Application US/10047542  
; Publication No. US20020168367A1  
; GENERAL INFORMATION:  
; APPLICANT: LARRICK, JAMES W.  
; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL  
; FILE REFERENCE: 030905.0004.CIP1  
; CURRENT APPLICATION NUMBER: US/10/047,542  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: PCT/US01/13932  
; PRIOR FILING DATE: 2001-04-28  
; PRIOR APPLICATION NUMBER: 60/200,298  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 101  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-047-542-16

Query Match 99.5%; Score 1889; DB 13; Length 353;  
Best Local Similarity 99.4%; Pred. No. 9.8e-117;  
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60  
Db 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60  
Qy 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYNPSQDVTVPCEVPSPTPTPTPTPTPTPTSP 120  
Db 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYNPSQDVTVPCEVPSPTPTPTPTPTPTPTSP 120  
Qy 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 180  
Db 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 180  
Qy 181 GCYSVSVSLGCAEPWNHKGFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Db 181 GCYSVSVSLGCAEPWNHKGFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Qy 241 ALNELVTLTCLARGFSPKDVLRWLQSQELPREKYLWTASRQPSQGTTFFAVTSILRV 300  
Db 241 ALNELVTLTCLARGFSPKDVLRWLQSQELPREKYLWTASRQPSQGTTFFAVTSILRV 300  
Qy 301 AAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 301 AAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353

RESULT 7  
US-10-072-012-797  
; Sequence 797, Application US/10072012

```
/ Publication No. US20040033493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zerhusen, Bryan
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangolli, Esha
/ APPLICANT: Padigaru, Muraidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Colman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grosse, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 797
/ LENGTH: 384
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-072-012-797

Query Match          99.5%; Score 1889; DB 15; Length 384;
Best Local Similarity 99.4%; Pred. No. 1.le-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 60
Db      32  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 91
QY      61  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
Db      92  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 151
QY     121  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 180
Db     152  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 211

Query Match          99.5%; Score 1889; DB 14; Length 393;
Best Local Similarity 99.4%; Pred. No. 1.le-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 60
Db      41  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 100
QY      61  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
Db     101  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 160
QY     121  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 180
Db     161  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 220
QY     181  GCVSVSSVLSGCAEPNHNKGTFTCTAAVPESKTPLTATLSKSGNTRPRPEVHLLPPSEEL 240
Db     221  GCVSVSSVLSGCAEPNHNKGTFTCTAAVPESKTPLTATLSKSGNTRPRPEVHLLPPSEEL 280
QY     241  ALNELVTTLTCLARGSPKDVLRVWLGSGOELPREKYLTVASRQEPSQGTTFITFAVTSILRV 300
Db     281  ALNELVTTLTCLARGSPKDVLRVWLGSGOELPREKYLTVASRQEPSQGTTFITFAVTSILRV 340
QY     301  AAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPTHVNVSVMVAEVDGTCY 353
Db     341  AAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPTHVNVSVMVAEVDGTCY 393

RESULT 9
US-10-104-047-3243

Query Match          99.5%; Score 1889; DB 15; Length 384;
Best Local Similarity 99.4%; Pred. No. 1.le-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 60
Db      32  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 91
QY      61  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
Db      92  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 151
QY     121  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 180
Db     152  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 211

Query Match          99.5%; Score 1889; DB 14; Length 393;
Best Local Similarity 99.4%; Pred. No. 1.le-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 60
Db      41  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 100
QY      61  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
Db     101  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 160
QY     121  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 180
Db     161  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 220
QY     181  GCVSVSSVLSGCAEPNHNKGTFTCTAAVPESKTPLTATLSKSGNTRPRPEVHLLPPSEEL 240
Db     221  GCVSVSSVLSGCAEPNHNKGTFTCTAAVPESKTPLTATLSKSGNTRPRPEVHLLPPSEEL 280
QY     241  ALNELVTTLTCLARGSPKDVLRVWLGSGOELPREKYLTVASRQEPSQGTTFITFAVTSILRV 300
Db     281  ALNELVTTLTCLARGSPKDVLRVWLGSGOELPREKYLTVASRQEPSQGTTFITFAVTSILRV 340
QY     301  AAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPTHVNVSVMVAEVDGTCY 353
Db     341  AAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPTHVNVSVMVAEVDGTCY 393

RESULT 8
US-10-221-945-3
/ Sequence 3, Application US/10221945
/ Publication No. US20030119737A1
/ GENERAL INFORMATION:
/ APPLICANT: INCYTE GENOMICS, INC.
/ APPLICANT: YUE, Henry
/ APPLICANT: HILLMAN, Jennifer L.
/ APPLICANT: BAUGHN, Mariah R.
/ TITLE OF INVENTION: HUMAN IMMUNE RESPONSE PROTEINS
/ FILE REFERENCE: PF-0765 PCT
/ CURRENT APPLICATION NUMBER: US/10/221,945
/ CURRENT FILING DATE: 2002-09-16
/ PRIOR APPLICATION NUMBER: PCT/US01/08518
/ PRIOR FILING DATE: 2001-03-15
/ PRIOR APPLICATION NUMBER: US 60/189,417
/ PRIOR FILING DATE: 2000-03-15
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PERL Program
/ SEQ ID NO 3
/ LENGTH: 393
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20030119737A1 2531065CDI
US-10-221-945-3

Query Match          99.5%; Score 1889; DB 14; Length 393;
Best Local Similarity 99.4%; Pred. No. 1.le-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 60
Db      41  ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPSPSODAS 100
QY      61  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
Db     101  GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 160
QY     121  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 180
Db     161  SCCHPRLSLHRLPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQPPDRDLC 220
QY     181  GCVSVSSVLSGCAEPNHNKGTFTCTAAVPESKTPLTATLSKSGNTRPRPEVHLLPPSEEL 240
Db     221  GCVSVSSVLSGCAEPNHNKGTFTCTAAVPESKTPLTATLSKSGNTRPRPEVHLLPPSEEL 280
QY     241  ALNELVTTLTCLARGSPKDVLRVWLGSGOELPREKYLTVASRQEPSQGTTFITFAVTSILRV 300
Db     281  ALNELVTTLTCLARGSPKDVLRVWLGSGOELPREKYLTVASRQEPSQGTTFITFAVTSILRV 340
QY     301  AAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPTHVNVSVMVAEVDGTCY 353
Db     341  AAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPTHVNVSVMVAEVDGTCY 393

RESULT 9
US-10-104-047-3243
```

```
; Sequence 3243, Application US/10104047
; Publication No. US200303236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US200303236392A1el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3243
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3243

Query Match      99.5%; Score 1889; DB 15; Length 491;
Best Local Similarity 99.4%; Pred. No. 1.4e-116;
Matches 351; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPOEPLSVTWSESGQGVTAARNFPPSQDAS 60
Db      139 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPOEPLSVTWSESGQGVTAARNFPPSQDAS 198

Qy      61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPSPSTPTPTSP 120
Db      199 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPSPSTPTPTSP 258

Qy      121 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 180
Db      259 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 318

Qy      181 GCYSVSVSVLSCGCAEPWNHNGKTFCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240
Db      319 GCYSVSVSVLPGCAEPWNHNGKTFCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 378

Qy      241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLITWASROQPSQGTTFFAVTSILRV 300
Db      379 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLITWASROQPSQGTTFFAVTSILRV 438

Qy      301 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353
Db      439 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 491

RESULT 11
US-10-108-260A-4290
; Sequence 4290, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: HI-A0106
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4290
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4290

Query Match      99.5%; Score 1889; DB 15; Length 491;
Best Local Similarity 99.4%; Pred. No. 1.4e-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPOEPLSVTWSESGQGVTAARNFPPSQDAS 60
Db      139 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPOEPLSVTWSESGQGVTAARNFPPSQDAS 198

Qy      61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPSPSTPTPTSP 120
Db      199 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPSPSTPTPTSP 258

Qy      121 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 180
Db      259 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 318

Qy      181 GCYSVSVSVLSCGCAEPWNHNGKTFCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240
Db      319 GCYSVSVSVLPGCAEPWNHNGKTFCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 378

Qy      241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLITWASROQPSQGTTFFAVTSILRV 300
Db      379 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLITWASROQPSQGTTFFAVTSILRV 438

Qy      301 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353
Db      439 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 491

RESULT 12
US-10-108-260A-4275
; Sequence 4275, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4262
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4262

Query Match      99.5%; Score 1889; DB 15; Length 491;
Best Local Similarity 99.4%; Pred. No. 1.4e-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPOEPLSVTWSESGQGVTAARNFPPSQDAS 60
Db      139 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPOEPLSVTWSESGQGVTAARNFPPSQDAS 198
```

```
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4275
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4275

Query Match          99.5%; Score 1889; DB 15; Length 494;
Best Local Similarity 99.4%; Pred No. 1.4e-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESGOGVTARNFPPSQDAS 60
Db 142 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESGOGVTARNFPPSQDAS 201
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTSP 120
Db 202 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTSP 261
QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTTTPSSGKSAVOGPPDRDLC 180
Db 262 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTTTPSSGKSAVOGPPDRDLC 321
QY 181 GCYSVSSVLSCGCAEPWNHKGKFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240
Db 322 GCYSVSSVLPGCAEPWNHKGKFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 381
QY 241 ALNELVTLTCLARGSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFATVTSILRV 300
Db 382 ALNELVTLTCLARGSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFATVTSILRV 441
QY 301 AAEWDKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353
Db 442 AAEWDKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 494

RESULT 13
US-10-072-012-798
; Sequence 798, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
```

```
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 798
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-798

Query Match          99.5%; Score 1889; DB 15; Length 494;
Best Local Similarity 99.4%; Pred No. 1.4e-116;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESGOGVTARNFPPSQDAS 60
Db 142 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESGOGVTARNFPPSQDAS 201
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTSP 120
Db 202 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTSP 261
QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTTTPSSGKSAVOGPPDRDLC 180
Db 262 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTTTPSSGKSAVOGPPDRDLC 321
QY 181 GCYSVSSVLSCGCAEPWNHKGKFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240
Db 322 GCYSVSSVLPGCAEPWNHKGKFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 381
QY 241 ALNELVTLTCLARGSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFATVTSILRV 300
Db 382 ALNELVTLTCLARGSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFATVTSILRV 441
QY 301 AAEWDKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353
Db 442 AAEWDKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 494

RESULT 14
US-10-108-260A-4085
; Sequence 4085, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4085
; LENGTH: 495
```

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-4085  
Query Match 99.5%; Score 1889; DB 15; Length 495;  
Best Local Similarity 99.4%; Pred. No. 1.4e-116;  
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60  
Db 143 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 202  
Qy 61 GDLYTTSQTLTPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTPTPTPTPTPTPT 120  
Db 203 GDLYTTSQTLTPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTPTPTPTPTPTPT 262  
Qy 121 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPTSSGKSAVQGPDRDLC 180  
Db 263 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPTSSGKSAVQGPDRDLC 322  
Qy 181 GCYSVSVSLSCAEPWNHGKFTTCTAAYPESKTELTATLSKSGNTFRPEVHLLPPPSSEEL 240  
Db 323 GCYSVSVSLPCCAEPWNHGKFTTCTAAYPESKTELTATLSKSGNTFRPEVHLLPPPSSEEL 382  
Qy 241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQSPSGTTFFAVTSILRV 300  
Db 383 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQSPSGTTFFAVTSILRV 442  
Qy 301 AAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 443 AAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 495

Search completed: October 25, 2005, 11:27:11  
Job time : 95.7482 secs

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-4114  
Query Match 99.5%; Score 1889; DB 15; Length 495;  
Best Local Similarity 99.4%; Pred. No. 1.4e-116;  
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60  
Db 143 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 202  
Qy 61 GDLYTTSQTLTPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTPTPTPTPTPTPT 120  
Db 203 GDLYTTSQTLTPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTPTPTPTPTPTPT 262  
Qy 121 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPTSSGKSAVQGPDRDLC 180  
Db 263 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPTSSGKSAVQGPDRDLC 322  
Qy 181 GCYSVSVSLSCAEPWNHGKFTTCTAAYPESKTELTATLSKSGNTFRPEVHLLPPPSSEEL 240  
Db 323 GCYSVSVSLPCCAEPWNHGKFTTCTAAYPESKTELTATLSKSGNTFRPEVHLLPPPSSEEL 382  
Qy 241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQSPSGTTFFAVTSILRV 300  
Db 383 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQSPSGTTFFAVTSILRV 442  
Qy 301 AAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 443 AAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 495

RESULT 15  
US-10-108-260A-4114  
; Sequence 4114, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: HI-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4114  
; LENGTH: 495  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-4114

Query Match 99.5%; Score 1889; DB 15; Length 495;  
Best Local Similarity 99.4%; Pred. No. 1.4e-116;  
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60  
Db 143 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 202  
Qy 61 GDLYTTSQTLTPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTPTPTPTPTPTPT 120  
Db 203 GDLYTTSQTLTPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPTPTPTPTPTPTPTPT 262  
Qy 121 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPTSSGKSAVQGPDRDLC 180  
Db 263 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPTSSGKSAVQGPDRDLC 322  
Qy 181 GCYSVSVSLSCAEPWNHGKFTTCTAAYPESKTELTATLSKSGNTFRPEVHLLPPPSSEEL 240  
Db 323 GCYSVSVSLPCCAEPWNHGKFTTCTAAYPESKTELTATLSKSGNTFRPEVHLLPPPSSEEL 382  
Qy 241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQSPSGTTFFAVTSILRV 300

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 10:53:52 ; Search time 57.8551 Seconds  
(without alignments)  
2359.799 Million cell updates/sec

Title: US-10-644-256-3\_COPY\_137\_489

Perfect score: 1898

Sequence: 1 ASPTSKVPLSLCSTQPDG.....GKPTHNVSVMAEVDGTCY 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	99.7	492	8	ADRI0249 Human pro
2	1893	99.7	494	7	Adm05393 Human pro
3	1893	99.7	494	8	Adq65690 Novel hum
4	1893	99.7	495	4	Asg62159 Human gen
5	1893	99.7	495	5	Abg63555 Human alb
6	1893	99.7	495	8	Adl76820 Albumin f
7	1893	99.7	496	8	Adr08452 Human pro
8	1893	99.7	497	7	Adb65619 Human pro
9	1893	99.7	497	8	Adq67021 Novel hum
10	1893	99.7	508	7	Adm04343 Human pro
11	1889	99.5	353	5	Aam47854 Human Ig-
12	1889	99.5	353	7	Ade97347 Human Iga
13	1889	99.5	353	8	Adnl1995 Immunoglo
14	1889	99.5	384	3	Aay88483 Cancer su
15	1889	99.5	384	5	Adil17261 Human NOV
16	1889	99.5	393	4	Aab82914 Human imm
17	1889	99.5	491	7	Adb65089 Human pro
18	1889	99.5	491	7	Adm05605 Human pro
19	1889	99.5	491	7	Adm05577 Human pro
20	1889	99.5	492	8	Adri10164 Human pro
21	1889	99.5	493	4	Aam93283 Human pol
22	1889	99.5	493	8	Adl30731 Human pro
23	1889	99.5	494	4	Aab95696 Human pro
24	1889	99.5	494	5	Adil17262 Human NOV
25	1889	99.5	494	7	Adm05590 Human pro

26	1889	99.5	495	7	Adm05400 Human pro
27	1889	99.5	495	7	Adm05429 Human pro
28	1889	99.5	496	7	Adb64852 Human pro
29	1889	99.5	496	7	Adm05373 Human pro
30	1889	99.5	497	3	Aay44723 Human imm
31	1889	99.5	497	8	Adp73848 Human ant
32	1889	99.5	499	8	Adq65976 Novel hum
33	1889	99.5	499	8	Adri10340 Human pro
34	1889	99.5	500	7	Adm05999 Human pro
35	1889	99.5	500	8	Adr09662 Human pro
36	1889	99.5	502	7	Adm05560 Human pro
37	1889	99.5	507	3	Aay96304 Human IGF
38	1889	99.5	508	7	Adb65079 Human pro
39	1889	99.5	630	7	Adf69023 Chloropla
40	1889	99.5	639	7	Adf68991 Chloropla
41	1885	99.3	495	7	Adm05592 Human pro
42	1885	99.3	497	7	Adm05559 Human pro
43	1882	99.2	500	7	Adm05570 Human pro
44	1881	99.1	497	8	Adq66283 Novel hum
45	1869	98.5	496	8	Adq65888 Novel hum

## ALIGNMENTS

## RESULT 1

ADRI0249  
ID ADR10249 standard; protein; 492 AA.

XX AC ADR10249;

XX DT 04-NOV-2004 (first entry)

XX DE Human protein useful for treating neurological disease Seq 3755.

XX KW human; oligo-capping method; diagnostic marker; gene therapy;

XX KW osteoporosis; neurological disease; Alzheimer's disease;

XX KW Parkinson's disease; dementia; short memory; cancer;

XX KW sense or motor function; emotional reaction; fear response; panic;

XX KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;

XX KW tranquiliser.

XX OS Homo sapiens.

XX PN EP1447413-A2.

XX PD 18-AUG-2004.

XX PF 12-FEB-2004; 2004EP-00003145.

XX PR 14-FEB-2003; 2003JP-00102207.

XX PR 09-MAY-2003; 2003JP-00131452.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;

XX PI Wakamatsu A, Ishii S, Nagai K, Irie R;

XX DR WPI; 2004-583265/57.

XX DR N-PSDB; ADR08293.

XX DR New 1995 cDNA, useful for treating osteoporosis, neurological diseases,

XX DR Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX PS Claim 1; SEQ ID NO 3755; 2686pp; English.

XX CC This invention relates to novel, isolated full length human cDNA

XX CC molecules and the encoded proteins thereof. Specifically, it refers to

XX CC cDNA clones obtained by an oligo-capping method, where none of these

XX CC clones are identical to any known human mRNAs. The present invention

XX CC describes an immunoassay to identify agonists and antagonists, as well as

XX CC antibodies, antisense molecules and siRNAs that can all be used to bind

XX CC to and modulate expression of the cDNA molecules. As such, these

CC molecules are useful for diagnostic markers or therapeutic targets for  
CC the various diseases or morbid states. In particular, they are useful in  
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
CC disease, Parkinson's disease, dementia, short memory and various cancers,  
CC as well as for maintaining equilibrium of sense or motor function, and  
CC for treating emotional reaction, fear response and panic. Accordingly,  
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
CC cytotatic and tranquiliser activities. This polypeptide is a protein  
CC encoded by a full length human cDNA sequence of the invention. NOTE: This  
CC sequence is not given in the sequence listing of the invention. NOTE: This  
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-  
CC office.  
XX  
SQ

Query Match 99.7%; Score 1893; DB 8; Length 492;  
Best Local Similarity 99.7%; Pred. No. 2.9e-119;  
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
SQ Sequence 492 AA;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESGOGVTARNPPPSQDAS 60  
DB 140 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESGOGVTARNPPPSQDAS 199  
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPTPTPTPTPTPTSP 120  
DB 200 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPTPTPTPTPTPTSP 259  
QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLDASGVFTTWPSSGKSAVQGPDRDL 180  
DB 260 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLDASGVFTTWPSSGKSAVQGPDRDL 319  
QY 181 GCYSVSVSLSGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240  
DB 320 GCYSVSVSLPGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 379  
QY 241 ALNELVTLTCLARGFSPKDVLRWLQSGQELPREKYLTVASRQEPQSGTTTFAVTSILRV 300  
DB 380 ALNELVTLTCLARGFSPKDVLRWLQSGQELPREKYLTVASRQEPQSGTTTFAVTSILRV 439  
QY 301 AAEDWKKGDTFSCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTCY 353  
DB 440 AAEDWKKGDTFSCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTCY 492

RESULT 2  
ADM05393  
ID ADM05393 standard; protein; 494 AA.

XX ADM05393;  
XX  
XX 20-MAY-2004 (first entry)  
XX  
XX Human protein of the invention SEQ ID NO:4078.  
XX  
XX human; gene therapy; diagnostic marker; pharmaceutical.  
XX  
XX Homo sapiens.  
XX

EP1347046-A1.

24-SEP-2003.

12-APR-2002; 2002EP-00008400.

22-MAR-2002; 2002JP-00137785.

(REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Negai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
XX WPI; 2003-723558/69.

DR N-PSDB; ADM02950.  
XX  
XX New polynucleotides and polypeptides are useful in gene therapy, for  
PT developing a diagnostic marker or medicines for regulating their  
PT expression and activity, or as a target of gene therapy.  
XX  
XX Claim 1; SEQ ID NO 4078; 305pp; English.  
XX  
XX The invention relates to a novel human polynucleotide and the encoded  
CC polypeptide. A polynucleotide of the invention may have a use in gene  
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
CC as a primer for synthesizing the polynucleotide or as a probe for  
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
CC useful in gene therapy, for developing a diagnostic marker or medicines  
CC for regulating their expression and activity, or as a target of gene  
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
CC are useful as pharmaceutical agents. The present sequence represents a  
CC protein sequence of the invention.  
XX  
SQ Sequence 494 AA;

Query Match 99.7%; Score 1893; DB 7; Length 494;  
Best Local Similarity 99.7%; Pred. No. 2.9e-119;  
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
SQ Sequence 494 AA;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESGOGVTARNPPPSQDAS 60  
DB 142 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPFLSVTWSESGOGVTARNPPPSQDAS 201  
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPTPTPTPTPTPTSP 120  
DB 202 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPTPTPTPTPTPTSP 261  
QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLDASGVFTTWPSSGKSAVQGPDRDL 180  
DB 262 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLDASGVFTTWPSSGKSAVQGPDRDL 321  
QY 181 GCYSVSVSLSGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240  
DB 322 GCYSVSVSLPGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 381  
QY 241 ALNELVTLTCLARGFSPKDVLRWLQSGQELPREKYLTVASRQEPQSGTTTFAVTSILRV 300  
DB 382 ALNELVTLTCLARGFSPKDVLRWLQSGQELPREKYLTVASRQEPQSGTTTFAVTSILRV 441  
QY 301 AAEDWKKGDTFSCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTCY 353  
DB 442 AAEDWKKGDTFSCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTCY 494

RESULT 3  
ADQ65690  
ID ADQ65690 standard; protein; 494 AA.

XX ADQ65690;

DT 07-OCT-2004 (first entry)

XX Novel human protein sequence #663.

XX osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
KW gene therapy; diagnostic marker; morbid state; osteoporosis;  
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;  
KW cancer.  
XX

OS Homo sapiens.

XX EP1440981-A2.

XX 28-JUL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX

```
PR 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakanatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Nagai K, Irie R;
XX WPI; 2004-535376/52.
XX N-PSDB; ADQ63502.
XX
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 2851; 2449pp; English.
XX
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
XX polypeptides, sequences hybridizing to these nucleotides, sequences
XX encoding partial polypeptides and sequences having 70% or 90% identity to
XX the nucleotide and protein sequences. The nucleotides and polypeptides
XX are useful as diagnostic markers or therapeutic target for the diseases
XX or morbid states. They are also useful for treating osteoporosis,
XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,
XX dementia and various cancers. This sequence corresponds to a protein
XX sequence of the invention.
XX
XX Sequence 494 AA;
XX
Query Match          99.7%; Score 1893; DB 8; Length 494;
Best Local Similarity 99.7%; Pred. No. 2.9e-119;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60
DB 142 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 201
QY 61 GDLVTTSSQLTLPATQCLAGKSVTCHVKHNTNPSQDVTVPVSTPTPTSPPTPSP 120
DB 202 GDLVTTSSQLTLPATQCLAGKSVTCHVKHNTNPSQDVTVPVSTPTPTSPPTPSP 261
QY 121 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDL 180
DB 262 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDL 321
QY 181 GCYSVSVSLSGCABPNHGKFTTCTAAYPSKTPLTATLSKSGNTFPEVHLLPPSEEL 240
DB 322 GCYSVSVSLSGCABPNHGKFTTCTAAYPSKTPLTATLSKSGNTFPEVHLLPPSEEL 381
QY 241 ALNELVTLTCLARGFSPKDLVRWLOGSQBELPREKYLVTWASRQPSQGTTFFAVTSILRV 300
DB 382 ALNELVTLTCLARGFSPKDLVRWLOGSQBELPREKYLVTWASRQPSQGTTFFAVTSILRV 441
QY 301 AAEWDKKGDTFSCMVGHEALPLAFTOKTIDRLAGKPTHNVSVVMAEVDGTCY 353
DB 442 AAEWDKKGDTFSCMVGHEALPLAFTOKTIDRLAGKPTHNVSVVMAEVDGTCY 494
RESULT 4
AAG62159
XX
XX AAG62159 standard; protein; 495 AA.
XX
XX AAG62159;
XX
XX 18-JUL-2001 (first entry)
XX
XX Human gene 4-encoded secreted protein HTOI042, SEQ ID NO:95.
XX
XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
```

```
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder; tumour;
KW endocrine disorder; infection; wound healing; vulnerability; cell culture;
KW chemotaxis; food additive; expression; binding partner identification.
XX
XX Homo sapiens.
XX W0200132910-A2.
XX 10-MAY-2001.
XX
XX 25-OCT-2000; 2000WO-US029362.
XX
XX 29-OCT-1999; 99US-0162240P.
XX 30-JUN-2000; 2000US-0215131P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Birse CE, Ni J, Soppet DR;
XX WPI; 2001-335835/35.
XX N-PSDB; AAH19173.
XX
XX Novel 27 isolated human secreted proteins and polynucleotides encoding
XX them useful for treating, diagnosing, preventing Alzheimer's disease,
XX Parkinson's disease, AIDS, rheumatoid arthritis, asthma.
XX
XX Claim 11; Page 497-499; 594pp; English.
XX
XX AAH19170-AAH19250 represent cDNAs corresponding to 27 human secreted
XX protein genes, and AAG62156-AAG62235 represent the proteins they encode.
XX AAG62236-AAG62293 represent human secreted protein fragments. The genes
XX and their corresponding secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 52 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, haematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
XX allergies, neurological disorders (e.g., Alzheimer's disease,
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin
XX disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX cardiovascular disorders, angiogenic disorders, kidney disorders,
XX gastrointestinal disorders, pregnancy-related disorders, endocrine
XX disorders, and infections. The proteins can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues, to identify their
XX cognate ligands or binding partners, and in chemotaxis, and can be used
XX as a food additive or preservative to modify storage properties.
XX Antibodies specific for a protein of the invention can be used in
XX alleviating symptoms associated with the disorders mentioned above, and
XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
XX immunosorbent assay (ELISA). The present sequence represents a human
XX secreted protein of the invention
XX
XX Sequence 495 AA;
```

```
Query Match          99.7%; Score 1893; DB 4; Length 495;
Best Local Similarity 99.7%; Pred. No. 2.9e-119;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60
DB 143 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 202
QY 61 GDLVTTSSQLTLPATQCLAGKSVTCHVKHNTNPSQDVTVPVSTPTPTSPPTPSP 120
DB 203 GDLVTTSSQLTLPATQCLAGKSVTCHVKHNTNPSQDVTVPVSTPTPTSPPTPSP 262
```

Qy	121	SCCHPRLSLHRPALEDLLLGSEANULTCTILTGLRDASGVTTFTWTPSSGKSAVQGPDRDLIC	180
Db	263	SCCHPRLSLHRPALEDLLLGSEANULTCTILTGLRDASGVTTFTWTPSSGKSAVQGPDRDLIC	322
Qy	181	GCYSVSSVLSGCAEPWNHGKFTCTTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL	240
Db	323	GCYSVSSVLPGCAEPWNHGKFTCTTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL	382
Qy	241	ALNELVTLTCLARGSPKDVLRVLWGQSOELPREKYLTWASRQEPSQGTTFITFAVTSILRV	300
Db	383	ALNELVTLTCLARGSPKDVLRVLWGQSOELPREKYLTWASRQEPSQGTTFITFAVTSILRV	442
Qy	301	AAEDWKKGDTSCVMVGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY	353
Db	443	AAEDWKKGDTSCVMVGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY	495
RESULT 5			
ABG63555	ID	ABG63555 standard; protein; 495 AA.	
XX	AC	ABG63555;	
XX	AC	ABG63555;	
XX	AC	ABG63555;	
DT	27-AUG-2002	(first entry)	
XX			
XX		Human albumin fusion protein #230.	
XX			
KW		Albumin fusion protein; therapeutic protein X; human albumin; HA;	
KW		human serum albumin; HSA; cancer; reproductive disorder;	
KW		digestive disorder; immune disorder; endocrine disorder;	
KW		haematopoietic disorder; neural disorder; connective disorder;	
KW		cysticotic; antiinfertility; antiinflammatory; antiulcer;	
KW		immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;	
KW		neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;	
KW		osteopathic; antiarthritic.	
XX			
OS		Homo sapiens.	
OS		Synthetic.	
XX			
PN		WO200177137-A1.	
XX			
PD		18-OCT-2001.	
XX			
XX			
PF		12-APR-2001; 2001WO-US011988.	
XX			
XX			
PR		12-APR-2000; 2000US-0229358P.	
PR		25-APR-2000; 2000US-0199384P.	
PR		21-DEC-2000; 2000US-0256931P.	
XX			
XX			
PA		(HUMA-) HUMAN GENOME SCI INC.	
XX			
PI		Rosen CA, Haseltine WA;	
XX			
DR		WPI; 2002-010886/01.	
XX			
PT		New fusion protein for treating disease e.g. diabetes comprises an	
PT		albumin fused to a therapeutic protein.	
XX			
PS		Claim 1; Page 712-714; 2102pp; English.	
XX			
CC			
CC		The present invention relates to albumin fusion proteins comprising a	
CC		therapeutic protein X and human albumin (HA, also known as human serum	
CC		albumin, HSA). The proteins are useful for treating a disease or disorder	
CC		that may be modulated by therapeutic protein X. The albumin extends the	
CC		shelf-life of protein X, and may increase its biological in vitro/in vivo	
CC		activity. The protein is useful for treating and diagnosing disorders	
CC		such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's	
CC		disease, ulcerative colitis), immune disorders (e.g. acquired	
CC		immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),	
CC		haematopoietic disorders, neural disorders (e.g. Alzheimer's,	
CC		Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,	
CC		schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).	
QC		ABG63326-ABG65518 represent albumin fusion proteins of the invention	

[illegible]

PS Disclosure; SEQ ID NO 302; 279pp; English.

XX The invention relates to a novel albumin fusion protein. The invention further relates to: a composition comprising the albumin fusion protein and a pharmaceutical carrier; a kit comprising the composition of the albumin fusion protein formula; a method of treating a disease or disorder in a patient comprising the step of administering the albumin fusion protein; a method of treating a patient with a disease or disorder that is modulated by therapeutic protein; X, or its fragment or variant; a method of extending the shelf life of therapeutic protein; X, or its fragment or variant; a nucleic acid molecule comprising a polynucleotide sequence encoding the albumin fusion protein; a vector comprising the nucleic acid molecule of the albumin fusion protein; and a host cell comprising the nucleic acid molecule of the albumin fusion protein. The albumin fusion protein and its compositions have the following activities: cytostatic, antineoplastic, antiarthritic, antiasthmatic, anti-HIV, immunosuppressive, antiinflammatory, antipsoriatic, antibacterial, osteopathic, dermatological, antitumor, immunomodulator, antiarrhythmic, cardiant, neurotropic, antiparkinsonian, nephrotropic, uropathic, neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic, hypertensive, and vulnerary. The albumin fusion protein nucleic acid may be used in gene therapy to treat disorders. The albumin fusion protein is useful for diagnosing, treating, preventing or ameliorating diseases or disorders comprising indication: Y. The diseases or disorders include: cancer (e.g. leukemia, colon, bone, breast, liver or lung cancer), cancer immune or haematopoietic diseases (e.g. anaemia, Hodgkin's disease, acute lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS, autoimmune disease, inflammatory bowel disease, psoriasis or Lyme disease), reproductive system disorders (e.g. prostatitis, inguinal hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-leydig tumours), musculoskeletal diseases (e.g. giant cell tumours, Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy or cachexia), cardiovascular disease (e.g. rhabdomyomas, heart disease, arrhythmia, cardiac arrest, heat valve disease, hypernatraemia or hyponatraemia), mixed foetal diseases (e.g. foetal alcohol syndrome, Down's syndrome, Patau syndrome, Turner's syndrome, Apert syndrome or Tay Sachs disease), excretory diseases (e.g. urinary incontinence, urinary tract infections or renal disorders), neural or sensory disease (e.g. Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis, cerebellar ataxia, attention deficit disorder, autism or obsessive compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or occupational lung disease), endocrine diseases (e.g. diabetes, Addison's disease or glomerulonephritis), digestive diseases (e.g. portal hypertension, irritable bowel disease, gastric atrophy or pancreatitis) or connective tissue or epithelial diseases (e.g. Crohn's disease, scleroderma, wound healing or epidermolysis bullosa). This sequence represents a therapeutic protein X relating to the albumin fusion protein of the invention. The sequence listing data for this specification was downloaded from the USPTO website.

XX

SQ Sequence 495 AA;

Query Match 99.7%; Score 1893; DB 8; Length 495;  
Best Local Similarity 99.7%; Pred. No. 2.9e-119; Indels 0; Gaps 0;  
Matches 352; Conservative 0; Mismatches 1;

QY 1 ASPTSPKVFPLSLGCTPDGNNVIACLVQGFPPQELPSVTWSESGQGVARNFPSPQDAS 60  
DB 143 ASPTSPKVFPLSLGCTPDGNNVIACLVQGFPPQELPSVTWSESGQGVARNFPSPQDAS 202

QY 61 GDLYTTSQTLPLATQCLAGKSVTHKHYTNPSQDVTVCPPVSTPTPTSPPTPTSP 120  
DB 203 GDLYTTSQTLPLATQCLAGKSVTHKHYTNPSQDVTVCPPVSTPTPTSPPTPTSP 262

QY 121 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLCL 180  
DB 263 SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLCL 322

QY 181 GCYSVSVSLSCAEPWNHGKFTTCTTAAYPESKKTTLATLSKSGNTFRPEVHLPPPPSEEL 240  
DB 323 GCYSVSVSLPCAEFPWNHGKFTTCTTAAYPESKKTTLATLSKSGNTFRPEVHLPPPPSEEL 382

QY 241 ALNELVTLTCLARGFSKDVLRWLQSGQELPREKYLTWASRQPSOGTTFITFAVTSILRV 300

DB 383 ALNELVTLTCLARGFSKDVLRWLQSGQELPREKYLTWASRQPSOGTTFITFAVTSILRV 442  
QY 301 AAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLACKPHTHVNVVMAEVDGTCY 353  
DB 443 AAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLACKPHTHVNVVMAEVDGTCY 495

RESULT 7  
ADR08452  
ID ADR08452 standard; protein; 496 AA.  
XX  
AC ADR08452;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DB Human protein useful for treating neurological disease Seq 1958.  
XX  
KW human; oligo-capping method; diagnostic marker; gene therapy;  
KW osteoporosis; neurological disease; Alzheimer's disease;  
KW Parkinson's disease; dementia; short memory; cancer;  
KW sense or motor function; emotional reaction; fear response; panic;  
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
KW tranquiliser.  
XX  
OS Homo sapiens.  
XX  
PN EPI447413-A2.  
XX  
PD 18-AUG-2004.  
XX  
PF 12-FEB-2004; 2004EP-00003145.  
XX  
PR 14-FEB-2003; 2003JP-00102207.  
XX  
PR 09-MAY-2003; 2003JP-00131452.  
XX  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Ieogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otaki T;  
PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
DR MPI; 2004-583265/57.  
XX  
DR N-PSDB; ADR06496.  
XX  
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX  
XX Claim 1; SEQ ID NO 1958; 2686pp; English.  
XX  
XX This invention relates to novel, isolated full length human cDNA  
XX molecules and the encoded proteins thereof. Specifically, it refers to  
XX cDNA clones obtained by an oligo-capping method, where none of these  
XX clones are identical to any known human mRNAs. The present invention  
XX describes an immunoassay to identify agonists and antagonists, as well as  
XX antibodies, antisense molecules and siRNAs that can all be used to bind  
XX to and modulate expression of the cDNA molecules. As such, these  
XX molecules are useful for diagnostic markers or therapeutic targets for  
XX the various diseases or morbid states. In particular, they are useful in  
XX gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
XX disease, Parkinson's disease, dementia, short memory and various cancers,  
XX as well as for maintaining equilibrium of sense or motor function, and  
XX for treating emotional reaction, fear response and panic. Accordingly,  
XX they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
XX cyostatic and tranquiliser activities. This polypeptide is a protein  
XX encoded by a full length human cDNA sequence of the invention. NOTE: This  
XX sequence is not given in the sequence listing of the specification but  
XX can be obtained on CD-ROM from the European Patent Office, Vienna Sub-  
XX office.  
XX  
SQ Sequence 496 AA;

Query Match 99.7%; Score 1893; DB 8; Length 496;  
Best Local Similarity 99.7%; Pred. No. 2.9e-119;

Matches	352;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	ASPTSPKVFPLSLCSTPDGNNVIACLVQGFPPQEPPLSVTWSESGQGV	TARNFPPSODAS	60					
Db	144	ASPTSPKVFPLSLCSTPDGNNVIACLVQGFPPQEPPLSVTWSESGQGV	TARNFPPSODAS	203					
Qy	61	GDLYTTSQSLTLPATQCLAGSKSVTCHVKHYTNPSODVTVP	CPVPSTPTTSPSP	120					
Db	204	GDLYTTSQSLTLPATQCLAGSKSVTCHVKHYTNPSQDVTVP	CPVPSTPTTSPSP	263					
Qy	121	SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTF	TWTPSSGKSAVGGPDRDLC	180					
Db	264	SCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTF	TWTPSSGKSAVGGPDRDLC	323					
Qy	181	GCYSVSSVLSGCAEPWNHGKFTCTAAYPESKTPLTATLSKSGNT	FRPEVHLLPPPESEL	240					
Db	324	GCYSVSSVLPCCAEPNWNGKFTCTAAYPESKTPLTATLSKSGNT	FRPEVHLLPPPESEL	383					
Qy	241	ALNELVTLTCLARGFSKDVILVRWLQSQEQLPREKYL	TWASRQPSQGTTF	300					
Db	384	ALNELVTLTCLARGFSKDVILVRWLQSQEQLPREKYL	TWASRQPSQGTTF	443					
Qy	301	AAEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNV	VMAEVDGTCY	353					
Db	444	AAEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNV	VMAEVDGTCY	496					
RESULT	8								
ADB65619									
ID	ADB65619	standard; protein; 497 AA.							
AC	ADB65619;								
XX									
XX									
DT	04-DEC-2003	(first entry)							
XX									
DE		Human protein encoded by clone THYMU20099060.							
XX									
KW		Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;							
KW		cell regeneration; membrane protein; signal transduction-related protein;							
KW		transcription-related protein; osteoporosis; neurological disease;							
KW		cancer; tumour.							
XX									
OS		Homo sapiens.							
XX									
PN		EPI308459-A2.							
XX									
PD		07-MAY-2003.							
XX									
PF		28-MAR-2002; 2002EP-00007401.							
XX									
PR		05-NOV-2001; 2001JP-00379298.							
PR		25-JAN-2002; 2002US-00350978.							
XX									
PA		(HELI-) HELIX RES INST.							
PA		(REAS-) RES ASSOC BIOTECHNOLOGY.							
XX									
PI		Isegami T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;							
PI		Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;							
PI		Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;							
XX									
XX		WPI; 2003-450961/43.							
DR		N-PSDB; ADB63649.							
XX									
PT		New polynucleotides and polypeptides, useful for developing a diagnostic							
PT		marker or medicines for regulation of their expression and activity, or							
PT		as targets of gene therapy.							
XX									
PS		Claim 1; Page: 222pp; English.							
XX									
CC		The invention discloses a polynucleotide comprising a sequence selected							
CC		from 1970 fully defined nucleotide sequences which encode novel							
CC		polypeptides. Also claimed is a polypeptide encoded by the polynucleotide							
CC		or its partial peptide, an antibody binding to the polypeptide or peptide							

of the polynucleotide, immunologically assaying the polypeptide or peptide of the polynucleotide by contacting the polypeptide or peptide with the antibody of the encoded protein, and observing the binding between the two, a transformant carrying the polynucleotide in an expressible manner and an antisense polynucleotide. The oligonucleotide is useful as a primer for synthesizing the polynucleotide, or as a probe for detecting the polynucleotide. The polynucleotides and encoded proteins are useful as pharmaceutical agents and many disease-related genes may be included in them, for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy. The genes are involved in tissue and/or cell regeneration. Membrane proteins, signal transduction-related proteins, transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours). The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a protein of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Query Match	99.7%	Score	1893	DB 7	Length	497			
Best Local Similarity	99.7%	Pred.	No. 3e-119						
Matches	352	Conservative	0	Mismatches	1	Indels	0	Gaps	
Qy	1	ASPTSPKVPFLSLCSTDPDGNVVIACLVQGRPPQPLSVTWSESQGVGTARNFPFSQDAS	60						
Db	145	ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQPLSVTWSESQGVGTARNFPFSQDAS	20						
Qy	61	GDLYTTSSQLTPATQCLAGKSVTCHVKHYTNPSQDVTVPQVPSTPTPSPTPTPTSP	12						
Db	205	GDLYTTSSQLTPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPSPTPTPTSP	26						
Qy	121	SCCHPRLSLHRPALBDLLLGSEANITCTLTGLRDASGVFTWTPSSGSKSAVOGPPDRDLC	18						
Db	265	SCCHPRLSLHRPALEDLLLGSEANITCTLTGLRDASGVFTWTPSSGSKSAVOGPPDRDLC	32						
Qy	181	GCYSVSSVLSGCAEPPNHHGKTFCTAAVPESKTPTATLSKSGNTFRPEVHLLPPPSSEL	24						
Db	325	GCYSVSSVLPGCAEPPNHHGKTFCTAAVPESKTPTATLSKSGNTFRPEVHLLPPPSSEL	38						
Qy	241	ALNELVTLTCLARGFSPKDLVRMLQSGQELPREKYLTWASRQEPSQGTTFVAVTSLRV	30						
Db	385	ALNELVTLTCLARGFSPKDLVRMLQSGQELPREKYLTWASRQEPSQGTTFVAVTSLRV	44						
Qy	301	AADWKKGDTFSCMGVGHREALPLAFQTKTIDRLAGKPTHNVNVSVMVAEVDGTCY	353						
Db	445	AADWKKGDTFSCMGVGHREALPLAFQTKIDRLAGKPTHNVNVSVMVAEVDGTCY	497						

RESULT 9	
ADQ67021	
ID	ADQ67021 standard; protein; 497 AA.
XX	
XX	ADQ67021;
XX	
XX	07-OCT-2004 (first entry)
DT	
XX	
XX	Novel human protein sequence #1994.
DE	
XX	osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW	gene therapy; diagnostic marker; morbid state; osteoporosis;
KW	neurological disease; Alzheimer's disease; Parkinson's disease; dementia
KW	cancer.
XX	
XX	
OS	Homo sapiens.
XX	
PN	EP1440981-A2.
XX	
XX	28-JUL-2004.
PD	
XX	



Human Ig-alpha1 heavy chain constant region amino acid sequence.

Human; immunoadhesin; intercellular adhesion molecule; ICAM-1;  
human rhinovirus; immunoglobulin heavy chain; J chain; HRV; common cold;  
transgenic plant.

Homo sapiens.

WO200183529-A2.

08-NOV-2001.

28-APR-2001; 2001WO-US013932.

28-APR-2000; 2000US-0200298P.

(PLAN-) PLANET BIOTECHNOLOGY INC.

Larrick JW, Wycoff KL;

WPI; 2002-041481/05.

N-PSDB; ABA05283.

Immunoadhesin for treating human rhinovirus infection comprises chimeric  
intercellular adhesion molecule-1, and optionally a J chain and secretory  
component in association.

Disclosure; Fig 7; 138pp; English.

The invention relates to an immunoadhesin comprising: (a) a chimeric  
intercellular adhesion molecule (ICAM)-1 comprising a rhinovirus receptor  
protein linked to at least a portion of an immunoglobulin heavy chain;  
and (b) optionally a J chain and secretory component associated with the  
chimeric ICAM-1 molecule. The immunoadhesin has plant-specific  
glycosylation and virucide activity. The immunoadhesin is useful for  
reducing infection by human rhinovirus (HRV) and hence the initiation or  
spread of the common cold by HRV. The immunoadhesin binds to HRV and  
reduces its infectivity, competing with cell surface ICAM-1 for binding  
sites, interfering with virus entry or uncoating and directing premature  
release of viral RNA and formation of empty capsids. Expression of the  
immunoadhesin in plants would be tetrameric, rather than dimeric.  
Immunoadhesin having multiple binding sites have a higher effective  
affinity for the virus, thereby increasing the effectiveness of the  
immunoadhesin. Association of secretory component and immunoglobulin J  
chain increases the stability of the immunoadhesin in the mucosal  
environment. Production is significantly less expensive in plants than in  
animal cell culture and production in plants is safer for human use,  
since plants are not known to harbor any animal viruses. The present  
sequence is that of a human immunoglobulin protein sequence, useful to  
the invention

XX SQ Sequence 353 AA;

Query Match 99.5%; Score 1889; DB 5; Length 353;  
Best Local Similarity 99.4%; Pred. No. 3.7e-119;  
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

1 ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

61 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTPSP 120

61 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTPSP 120

121 SCCHPRLSLHRPALEDLLLGSEANLACTLTGLRDASGVTFWTPTSSGKSAVQGPDPDRLC 180

121 SCCHPRLSLHRPALEDLLLGSEANLACTLTGLRDASGVTFWTPTSSGKSAVQGPDPDRLC 180

181 GCYSVSVSLSGCAEPWNHGKFTCTAAYPESKTPLTATLSKSGNTPRPEVHLLPPPSSEEL 240

181 GCYSVSVSLPGCAEPWNHGKFTCTAAYPESKTPLTATLSKSGNTPRPEVHLLPPPSSEEL 240

QY 241 ALNELVTLTCLARGSPKDVLRVLQGSQELPREKYLTVASRQEPSQGTTFVAVTSILRV 300  
Db 241 ALNELVTLTCLARGSPKDVLRVLQGSQELPREKYLTVASRQEPSQGTTFVAVTSILRV 300  
QY 301 AAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 301 AAEDWKKGDTFSCMVGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 353

RESULT 12

ADE97347

ID ADE97347 standard; protein; 353 AA.

XX ADE97347;

DT 12-FEB-2004 (first entry)

DE Human IgA1 heavy chain constant region protein - SEQ ID 16.

XX immunoadhesin; immunoglobulin heavy chain; J chain; joining; toxin;  
KW virucide; antibacterial; anthrax; rhinovirus infection; common cold;  
KW intercellular adhesion molecule; ICAM-1; human; constant region; IGA.

XX Homo sapiens.

XX WO2003064992-A2.

XX 07-AUG-2003.

PF 25-OCT-2002; 2002WO-US034197.

PR 26-OCT-2001; 2001US-00047542.

XX (PLAN-) PLANET BIOTECHNOLOGY INC.

PA (LARR/) LARRICK J W.

PA (WYCO/) WYCOFF K L.

XX Larrick JW, Wycoff KL;

XX WPI; 2003-636816/60.

DR N-PSDB; ADE97346, ADE97374.

XX New immunoadhesin, useful for treating anthrax and rhinovirus, comprises  
chimeric toxin receptor protein linked to immunoglobulin heavy chain, and  
J chain and secretory component associated with the chimeric toxin  
receptor protein.

XX Disclosure; SEQ ID NO 16; 288pp; English.

XX The invention relates to a novel immunoadhesin comprising a chimeric  
toxin receptor protein consisting of a toxin receptor protein linked to  
at least a portion of an immunoglobulin heavy chain with a J (joining)  
chain and secretory component (SC) associated with the chimeric toxin  
receptor protein. The immunoadhesin comprises a chimeric bacterial or  
viral toxin receptor protein and the immunoadhesin has plant-specific  
glycosylation. The immunoadhesin of the invention demonstrates virucide  
and antibacterial activities and may be useful for reducing the binding  
of a viral or bacterial antigen to a host cell and thus for treating or  
preventing anthrax, as well as human rhinovirus infection which results  
in the common cold. The current sequence is that of the human  
immunoadhesin-related protein of the invention.

XX SQ Sequence 353 AA;

Query Match 99.5%; Score 1889; DB 7; Length 353;  
Best Local Similarity 99.4%; Pred. No. 3.7e-119;  
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

1 ASPTSPKVPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

QY 61 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTPSP 120

Db 61 GDLYTSSQLTPATQCLAGSVTCHVKHYTNPSQDVTVPVPSPTPTPSP 120  
QY 121 SCCHPRLSLHRPALEDLLGLSEANLCTLTGLRDASGVTFWTWPSGKSAVQGPDRDLC 180  
Db 121 SCCHPRLSLHRPALEDLLGLSEANLCTLTGLRDASGVTFWTWPSGKSAVQGPDRDLC 180  
QY 181 GCYSVSVSLGCAEPWNHKGKTFCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240  
Db 181 GCYSVSVSLGCAEPWNHKGKTFCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240  
QY 241 ALNELVTLTCLARGFSPKDVLRWLOGSQELPREKYLTVASROEPSOGTTTFAVTSILRV 300  
Db 241 ALNELVTLTCLARGFSPKDVLRWLOGSQELPREKYLTVASROEPSOGTTTFAVTSILRV 300  
QY 301 AABDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 301 AABDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353

## RESULT 13

ADN11995  
ID ADN11995 standard; protein; 353 AA.

XX  
AC ADN11995;

XX 17-JUN-2004 (first entry)

XX Immunoglobulin A heavy chain.

XX immunoglobulin A polypeptide; diabetic retinopathy;

XX immunoglobulin A heavy chain.

XX Homo sapiens.

XX WO2004007554-A1.

XX 22-JAN-2004.

XX 20-MAR-2003; 2003WO-KR000544.

XX 16-JUL-2002; 2002KR-00041771.

XX (EYEG-) EYEGENE INC.

XX Yoo WI, Lee SH, Park K, Cho YJ, Ahn BY, Kwon OW;

XX WPI; 2004-122895/12.

XX N-PSDB; ADN11997.

XX New immunoglobulin A polypeptide and encoding nucleic acid molecule,  
XX useful for diagnosing diabetic retinopathy using an immunologic method.

XX Claim 1; SEQ ID NO 1; 27pp; English.

XX The present invention relates to an immunoglobulin A polypeptide for

XX diagnosing diabetic retinopathy. The composition and method are useful

XX for diagnosing diabetic retinopathy. The present sequence represents

XX immunoglobulin A heavy chain.

XX Sequence 353 AA;

XX Query Match 99.5%; Score 1889; DB 8; Length 353;

XX Best Local Similarity 99.4%; Pred. No. 3.7e-119;

XX Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTVARNPPSQDAS 60

Db 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTVARNPPSQDAS 60

QY 61 GDLYTSSQLTPATQCLAGSVTCHVKHYTNPSQDVTVPVPSPTPTPSP 120

Db 61 GDLYTSSQLTPATQCLAGSVTCHVKHYTNPSQDVTVPVPSPTPTPSP 120

QY 121 SCCHPRLSLHRPALEDLLGLSEANLCTLTGLRDASGVTFWTWPSGKSAVQGPDRDLC 180  
Db 121 SCCHPRLSLHRPALEDLLGLSEANLCTLTGLRDASGVTFWTWPSGKSAVQGPDRDLC 180  
QY 181 GCYSVSVSLGCAEPWNHKGKTFCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240  
Db 181 GCYSVSVSLGCAEPWNHKGKTFCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240  
QY 241 ALNELVTLTCLARGFSPKDVLRWLOGSQELPREKYLTVASROEPSOGTTTFAVTSILRV 300  
Db 241 ALNELVTLTCLARGFSPKDVLRWLOGSQELPREKYLTVASROEPSOGTTTFAVTSILRV 300  
QY 301 AABDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 301 AABDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353

## RESULT 14

AA98483  
ID AA98483 standard; protein; 384 AA.

XX  
AC AA98483;

XX 03-AUG-2000 (first entry)

XX Cancer suppressor gene product.

XX Cancer suppressor gene; large intestine cancer; treatment; tumour;

XX chromosome 14q32.

XX Homo sapiens.

XX WO200018911-A1.

XX 06-APR-2000.

XX 30-SEP-1998; 98WO-CN000208.

XX 30-SEP-1998; 98WO-CN000208.

XX (ZHEN/) ZHENG S.

XX Zheng S, Cao J, Cao W, Geng L, Zhang Y;

XX WPI; 2000-293150/25.

XX N-PSDB; AAA15694.

XX Cancer-suppressor gene down-regulated in large intestine cancer, located  
XX in human chromosome 14q32, useful for diagnosis and treatment of tumors  
XX particularly large intestine cancer.

XX Claim 1; Fig 2; 42pp; Chinese.

XX This sequence represents the protein encoded by a cancer-suppressor gene,  
XX which is down regulated in large intestine cancer. The gene is located on  
XX chromosome 14q32. The invention also relates to a vector comprising the  
XX gene, a host cell transformed by the vector, and a process for culturing  
XX the host cell and recovering the expression product. The gene, encoded  
XX polypeptide and antibody are used in the diagnosis and treatment of  
XX cancers particularly large intestine cancer

XX Sequence 384 AA;

XX Query Match 99.5%; Score 1889; DB 3; Length 384;

XX Best Local Similarity 99.4%; Pred. No. 4.1e-119;

XX Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTVARNPPSQDAS 60

Db 32 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTVARNPPSQDAS 91

QY 61 GDLYTSSQLTPATQCLAGSVTCHVKHYTNPSQDVTVPVPSPTPTPSP 120

Db 92 GDLYTTSSQLTPATOCLAGKSVTCHVKHYTNPSQDVTVPSPPTPTPTSP 151  
QY 121 SCCHPRLSHRPALEDLLGSEANLTCTLTGLRDASGVFTTWTSSGKSAVGQPPDRDLC 180  
Db 152 SCCHPRLSHRPALEDLLGSEANLTCTLTGLRDASGVFTTWTSSGKSAVGQPPDRDLC 211  
QY 181 GCYSVSSVLGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTRPEVHLLPPSEEL 240  
Db 212 GCYSVSSVLGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTRPEVHLLPPSEEL 271  
QY 241 ALNELVTLTCLARGSPKDVLRWLQGSQELPREKYLITWASRQEPSQGTTFEAVTSILRV 300  
Db 272 ALNELVTLTCLARGSPKDVLRWLQGSQELPREKYLITWASRQEPSQGTTFEAVTSILRV 331  
QY 301 AAEDWKKGDTFCMGVGEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 332 AAEDWKKGDTFCMGVGEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 384

RESULT 15

AD117261

ID AD117261 standard; protein; 384 AA.

XX AC

XX AD117261;

XX DT 15-APR-2004 (first entry)

XX DE Human NOVX protein homologue SeqID 797.

XX KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;

XX KW inflammation; autoimmune disorder; allergy; blood disorder;

XX KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;

XX KW immunoglobulin (IgA nephropathy; cirrhosis; arthritis;

XX KW Alzheimer's disease; infection; str.

XX OS Homo sapiens.

XX PN WO200268649-A2.

XX PD 06-SEP-2002.

XX PF 31-JAN-2002; 2002WO-US002785.

XX PR 31-JAN-2001; 2001US-0265395P.

XX PR 31-JAN-2001; 2001US-0265412P.

XX PR 31-JAN-2001; 2001US-0265514P.

XX PR 02-FEB-2001; 2001US-0266406P.

XX PR 05-FEB-2001; 2001US-0266767P.

XX PR 07-FEB-2001; 2001US-0266975P.

XX PR 07-FEB-2001; 2001US-0267057P.

XX PR 08-FEB-2001; 2001US-0267459P.

XX PR 09-FEB-2001; 2001US-0267823P.

XX PR 15-FEB-2001; 2001US-0268974P.

XX PR 26-FEB-2001; 2001US-0271664P.

XX PR 27-FEB-2001; 2001US-0271839P.

XX PR 27-FEB-2001; 2001US-0271855P.

XX PR 02-MAR-2001; 2001US-0272788P.

XX PR 02-MAR-2001; 2001US-0273046P.

XX PR 14-MAR-2001; 2001US-0275925P.

XX PR 14-MAR-2001; 2001US-0275947P.

XX PR 14-MAR-2001; 2001US-0275950P.

XX PR 15-MAR-2001; 2001US-0276488P.

XX PR 15-MAR-2001; 2001US-0276450P.

XX PR 16-MAR-2001; 2001US-0276397P.

XX PR 20-MAR-2001; 2001US-0276768P.

XX PR 26-MAR-2001; 2001US-0278652P.

XX PR 26-MAR-2001; 2001US-0278775P.

XX PR 29-MAR-2001; 2001US-0279882P.

XX PR 29-MAR-2001; 2001US-0279884P.

PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282992P.  
PR 11-APR-2001; 2001US-0283083P.  
PR 23-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285749P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 29-MAY-2001; 2001US-0294047P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 08-JUN-2001; 2001US-0296964P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299324P.  
PR 13-AUG-2001; 2001US-0312020P.  
PR 16-AUG-2001; 2001US-0312889P.  
PR 16-AUG-2001; 2001US-0312908P.  
PR 21-AUG-2001; 2001US-0313390P.  
PR 28-AUG-2001; 2001US-0315470P.  
PR 31-AUG-2001; 2001US-0316447P.  
PR 07-SEP-2001; 2001US-0318115P.  
PR 07-SEP-2001; 2001US-0318118P.  
PR 12-SEP-2001; 2001US-0318740P.  
PR 19-SEP-2001; 2001US-0323379P.  
PR 18-OCT-2001; 2001US-0330245P.  
PR 18-OCT-2001; 2001US-0330308P.  
PR 14-NOV-2001; 2001US-0332701P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;  
PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;  
PI Gerlach VL, Taupier RJ, Gusev VV, Colman SD, Wolenc AP, Pena CE;  
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX WPI; 2002-706998/76.

XX New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
XX Disclosure; SEQ ID NO 797; 1498pp; English.

XX This invention relates to a novel nucleic acids, and encoded polypeptides  
XX thereof, which have properties related to the stimulation of biochemical  
XX or physiological responses in a cell, tissue, organ or organism.  
XX Specifically, it refers to the use of biologically active fragments for  
XX diagnostic and prognostic assays and furthermore in the treatment of  
XX diverse pathological conditions. The present invention describes novel  
XX human and murine NOVX proteins, as well as methods to modulate their  
XX expression using antisense oligos, ribozymes and peptide nucleic acids.  
XX The NOVX polypeptides, polynucleotides and antibodies are useful in  
XX treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,  
XX atherosclerosis, cancer and diabetes. Furthermore, they may be used in  
XX treating or preventing diseases such as inflammation, autoimmune  
XX disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
XX (AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis,  
XX arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
XX and epilepsy. Accordingly, these molecules have many activities including  
XX cytotstatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,  
XX haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,  
XX antiasthmatic, nephrotropic, antiarthritic, hepatotropic,  
XX neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
XX relaxant and anticonvulsant. In addition, they are useful in screening  
XX assays to identify small molecules that modulate or inhibit, for example,  
XX neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
XX used as in chromosome mapping, tissue typing, preventive medicine and  
XX pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
XX of the invention.

XX SQ Sequence 384 AA;

Query Match 99.5%; Score 1889; DB 5; Length 384;  
Best Local Similarity 99.4%; Pred. No. 4.1e-119;

Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
Qy 1	ASPTSPKVFPLSLCSTPDGNNWIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPSPQDAS 60
Db	
Qy 32	ASPTSPKVFPLSLCSTPDGNNWIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPSPQDAS 91
Db	
Qy 61	GDLYTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPQVPSTPPTPSPPTPSP 120
Db	
Qy 92	GDLYTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPQVPSTPPTPSPPTPSP 151
Db	
Qy 121	SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQGPDRDLC 180
Db	
Qy 152	SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTFTWTPSSGKSAVQGPDRDLC 211
Db	
Qy 181	GCYSVSSVLSGCABPWNHGKFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240
Db	
Qy 212	GCYSVSSVLSGCABPWNHGKFTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 271
Db	
Qy 241	ALNELVTLTCLARGFSPKDVLRWLOGSOELPREKYLTVASROEPSOGTTFAVTSILRV 300
Db	
Qy 272	ALNELVTLTCLARGFSPKDVLRWLOGSOELPREKYLTVASROEPSOGTTFAVTSILRV 331
Db	
Qy 301	AAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353
Db	
Qy 332	AAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 384
Db	

Search completed: October 25, 2005, 11:05:49  
Job time : 59.8551 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 10:59:42 ; Search time 12.158 Seconds  
(without alignments)  
2793.604 Million cell updates/sec

Title: US-10-644-256-3\_COPY\_137\_489

Perfect score: 1898  
Sequence: 1 ASPTSPKVFPLSLCSTQPDG.....GKPTHVNVSVMAEVDGTCY 353

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1889	99.5	353	1 A1HU	Ig alpha-1 chain C
2	1838	96.8	352	2 S05500	Ig mu chain C
3	1710.5	90.1	340	2 B22360	Ig alpha-2 chain C
4	1693.5	89.2	340	2 I56230	Ig alpha-2 chain -
5	1656.5	87.3	340	1 A2HU	Ig alpha-2 chain C
6	1275.5	67.2	342	2 I47175	Ig alpha chain C
7	1113	58.6	220	2 C22360	Ig alpha-2 chain C
8	1112.5	58.6	342	2 A45966	Ig alpha chain C
9	1066.5	56.2	344	1 AHMS	Ig alpha chain C
10	1035.5	54.6	357	2 S09267	Ig alpha chain C
11	1022	53.8	357	2 S09269	Ig alpha chain C
12	1016	53.5	357	2 S09268	Ig alpha chain C
13	986	51.9	358	2 S09274	Ig alpha chain C
14	984	51.8	347	2 S09270	Ig alpha chain C
15	980	51.6	348	2 S09270	Ig alpha chain C
16	975	51.4	348	2 S09273	Ig alpha chain C
17	970.5	51.1	339	2 S09264	Ig alpha chain C
18	959.5	50.6	343	2 S09272	Ig alpha chain C
19	949	50.0	338	2 S09276	Ig alpha chain C
20	945	49.8	352	2 S09266	Ig alpha chain C
21	943	49.7	360	2 S09271	Ig alpha chain C
22	893.5	46.5	335	2 S09275	Ig alpha chain C
23	838	44.2	299	1 AHRB	Ig alpha chain C
24	710.5	37.4	145	2 S03298	Ig alpha chain C
25	619	32.6	132	2 I61901	Ig alpha chain - o
26	556	29.3	585	2 A46507	Ig alpha chain - c
27	544	28.7	455	1 MHMS	Ig mu chain C regi
28	544	28.7	455	2 A24976	Ig mu chain C regi
29	538	28.3	453	2 S37768	Ig mu chain C regi

30	534	28.1	454	1	MHHY	Ig mu chain C regi
31	528.5	27.8	458	1	MHRB	Ig mu chain C regi
32	526.5	27.7	452	1	MHRU	Ig mu chain C regi
33	518.5	27.3	457	2	S03961	Ig mu chain C regi
34	518	27.3	592	2	S25705	Ig mu chain - shee
35	512	27.0	448	2	S03186	Ig heavy chain C r
36	501	26.4	391	1	MHHUBT	Ig mu heavy chain
37	490.5	25.8	454	2	A46532	Ig mu chain C regi
38	488	25.7	343	2	S25644	Ig mu chain C regi
39	475.5	25.1	476	1	MHMSM	Ig mu chain C regi
40	473.5	24.9	450	1	MHDG	Ig mu chain C regi
41	469.5	24.7	474	2	S15590	Ig heavy chain - h
42	469	24.7	627	2	S14683	Ig mu chain precu
43	466.5	24.6	479	1	MHRBM	Ig mu chain C regi
44	458	24.1	473	1	MHRUM	Ig mu chain C regi
45	445.5	23.5	328	2	I47161	Ig gamma 3 chain c

ALIGNMENTS

RESULT 1

A1HU

Ig alpha-1 chain C region - human

C;Species: Homo sapiens (man)

C;Date: 22-May-1981 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004

C;Accession: A22360; A92249; A91662; S38979; B53110; A02171

R;Flanagan, J.G.; Lefranc, M.P.; Rabbitts, T.H.

Cell 36; 681-688, 1984

A;Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1 and

A;Reference number: A94653; MUID:84130179; PMID:6421489

A;Accession: A22360

A;Molecule type: DNA

A;Residues: 1-353 <FLA>

A;Cross-references: UNIPROT:P01876

R;Putnam, F.W.; Liu, Y.S.V.; Low, T.L.K.

J. Biol. Chem. 254, 2863-2874, 1979

A;Title: Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal IgA1 protease

A;Reference number: A92249; MUID:79151016; PMID:107164

A;Contents: myeloma protein Bur; disulfide bonds

A;Accession: A92249

A;Molecule type: protein

A;Residues: 1-16, 'Z', '18', 'B', '20', 'B', '22-34', 'Q', '36-45', 'Z', '47-51', 'B', '53-56', 'ZB', '59-61', 'B', '63', '303', 'B', '305-346', 'Q', '348-353 <PUT>

A;Note: this is the final paper in a series

A;Note: amidation states of residues 178, 197, 238, 239, 243, 244, 287, and 318 were take

R;Kratz, H.; Altevogt, P.; Ruban, E.; Kortt, A.; Starosick, K.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 356, 1337-1342, 1975

A;Title: The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.), II: the am

A;Reference number: A91662; MUID:76023781; PMID:809331

A;Accession: A91662

A;Molecule type: protein

A;Residues: 1-16, 'Z', '18', 'B', '20', 'B', '22-34', 'Q', '36-45', 'Z', '47-51', 'B', '53-56', 'ZB', '59-61', 'B', '63', 'R', '232-237', 'QQ', '240-243', 'Q', '245-283', 'O', '285-289', 'E', '291-303', 'B', '305-353 <KRA>

A;Experimental source: Myeloma protein Tro

R;Fallgreen-Gebauer, E.; Gebauer, W.; Bastian, A.; Kratzin, H.D.; Eifert, H.; Zimmermann

Biol. Chem. Hoppe-Seyler 374, 1023-1028, 1993

A;Title: The covalent linkage of secretory component to IgA. Structure of sIgA.

A;Reference number: S38978; MUID:94121784; PMID:8292260

A;Accession: S38979

A;Molecule type: protein

A;Residues: 188-196, 'D', '198-201 <FAL>

R;Yang, C.Y.; Kratzin, H.; Gotz, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 360, 1919-1940, 1979

A;Title: Die Primaerstruktur eines monoklonalen IgA1-Immunglobulins (Myelomprotein Tro).

A;Reference number: A91684; MUID:80114124; PMID:393607

A;Contents: annotation; Tro; disulfide bonds

A;Note: Cys-14 bonds to a light chain

R;Calero, M.; Escribano, J.; Grubb, A.; Mendez, E.

J. Biol. Chem. 269, 384-389, 1994

A;Title: Location of a novel type of interpeptide chain linkage in the human protein

A;Reference number: A53110; MUID:94103241; PMID:7506257

A;Accession: B53110

A:Molecule type: protein

A:Residues: 346-351,'X',353 <CAL>

C:Genetics:

A:Gene: GDB:IGHAI

A:Cross-references: GDB:119332; OMIM:146900

A:Map position: 14q32.33-14q32.33

A:Introns: 1/1; 103/1; 223/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds.

C:Superfamily: Immunoglobulin C region; immunoglobulin homology

F:140-206/Domain: duplication; glycoprotein; heterotetramer; immunoglobulin; tryptophan

F:243-315/Domain: immunoglobulin homology <IM2>

F:26-85,77-101,123-180,147-204,250-313/Disulfide bonds: #status experimental

F:105,111,113,119,121/Binding site: carbohydrate (ser) (covalent) #status experimental

F:122,182/Disulfide bonds: interchain #status experimental

F:144,340/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:192/Disulfide bonds: interchain (to secretory component) (partial) #status experimental

F:192/Binding site: cysteine (Cys) (covalent) (partial) #status experimental

F:192/Disulfide bonds: interchain (partial) #status experimental

F:352/Cross-link: alpha-1-microglobulin-ig alpha complex chromophore (Cys) (interchain)

F:352/Disulfide bonds: interchain (to J chain) (partial) #status experimental

F:352/Disulfide bonds: interchain (partial) #status experimental

Query Match 99.5%; Score 1889; DB 1; Length 353;

Best Local Similarity 99.4%; Pred. No. 2.8e-106;

Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

DB 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

QY 61 GDLYTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPCVPSTPTPTSP 120

DB 61 GDLYTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPCVPSTPTPTSP 120

QY 121 SCCHPRLSLHRRPALEDLLLGSEANLTCTLTGLRDASGVTTWTPSSGKSAVQGPDRDLC 180

DB 121 SCCHPRLSLHRRPALEDLLLGSEANLTCTLTGLRDASGVTTWTPSSGKSAVQGPDRDLC 180

QY 181 GCVSVSVSLGCAEPNWHGKTFTCTAAYPESKTPLTATLSKSGNTPRPEVHLLPPSEEL 240

DB 181 GCVSVSVSLGCAEPNWHGKTFTCTAAYPESKTPLTATLSKSGNTPRPEVHLLPPSEEL 240

QY 241 ALNELVTLTCLARGSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFFAVTSILRV 300

DB 241 ALNELVTLTCLARGSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFFAVTSILRV 300

QY 301 AAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353

DB 301 AAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353

RESULT 2

S05500

Ig alpha-1 chain C region - gorilla (fragment)

C:Species: Gorilla gorilla (gorilla)

C>Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 23-Jul-1999

C:Accession: S05500

R:Kawamura, S.; Omoto, K.; Ueda, S.

Nucleic Acids Res. 17, 6732, 1989

A:Title: Nucleotide sequence of the gorilla immunoglobulin alpha 1 gene.

A:Reference number: S05500; MUID:89386006; PMID:2506527

A:Accession: S05500

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-352 <KAW>

A:Cross-references: EMBL:X15045; NID:g22900; PIDN:CAA33147.1; PID:g22901

C:Genetics:

A:Introns: 102/1; 222/1

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:242-314/Domain: immunoglobulin homology <IM2>

Query Match 96.8%; Score 1838; DB 2; Length 352;

Best Local Similarity 96.9%; Pred. No. 3.2e-103;

Matches 341; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 SPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 61

DB 1 SPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

QY 62 DLYTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPCVPSTPTPTSP 121

DB 61 DLYTSSQLTLPATOCPCDKSVTCHVHYTNPSQDVTVPVPCVPSTPTPTSP 120

QY 122 CCHPRLSLHRRPALEDLLLGSEANLTCTLTGLRDASGVTTWTPSSGKSAVQGPDRDLC 181

DB 121 CCHPRLSLHRRPALEDLLLGSEANLTCTLTGLRDASGVTTWTPSSGKSAVQGPDRDLC 180

QY 182 GCVSVSVSLGCAEPNWHGKTFTCTAAYPESKTPLTATLSKSGNTPRPEVHLLPPSEELA 241

DB 181 GCVSVSVLPGCAEPNWHGKTFTCTAAYPESKTPLTATLSKSGNMFAPEVHLLPPSEELA 240

QY 242 LNELVTLTCLARGSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFFAVTSILRV 301

DB 241 LNELVTLTCLARGSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFFAVTSILRV 300

QY 302 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353

DB 301 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 352

RESULT 3

B22360

Ig alpha-2 chain C region (allotype A2m(1)) - human

C:Species: Homo sapiens (man)

C>Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004

C:Accession: B22360

R:Planagan, J.G.; Lefranc, M.P.; Rabbitts, T.H.

Cell 36, 681-688, 1984

A:Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1 and

A:Reference number: A94653; MUID:84130179; PMID:6421489

A:Accession: B22360

A:Molecule type: DNA

A:Residues: 1-340 <FLA>

A:Cross-references: UNIPROT:P01877

C:Genetics:

A:Gene: GDB:IGHA2

A:Cross-references: GDB:119333; OMIM:147000

A:Map position: 14q32.33-14q32.33

A:Introns: 1/1 103/1; 210/1

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:230-302/Domain: immunoglobulin homology <IM2>

Query Match

Best Local Similarity 90.1%; Score 1710.5; DB 2; Length 340;

Matches 321; Conservative 7; Mismatches 12; Indels 13; Gaps 1;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

DB 1 ASPTSPKVPFLSLDSTPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60

QY 61 GDLYTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPCVPSTPTPTSP 120

DB 61 GDLYTSSQLTLPATOCPCDKSVTCHVHYTNPSQDVTVPVPCVPSTPTPTSP 107

QY 121 SCCHPRLSLHRRPALEDLLLGSEANLTCTLTGLRDASGVTTWTPSSGKSAVQGPDRDLC 180

DB 108 PCCHPRLSLHRRPALEDLLLGSEANLTCTLTGLRDASGVTTWTPSSGKSAVQGPDRDLC 167

QY 181 GCVSVSVSLGCAEPNWHGKTFTCTAAYPESKTPLTATLSKSGNTPRPEVHLLPPSEEL 240

DB 168 GCVSVSVLPGCAEPNWHGKTFTCTAAHPPELKTPLTITKSGNTRFPEVHLLPPSEEL 227



A;Cross-references: EMBL:U12594; NID:g555826; PIDN:AAA65943.1; PID:g555827  
C:Genetics:  
A:Gene: IgA2alpha  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:232-304/Domain: immunoglobulin homology <IMM>  
Query Match 67.2%; Score 1275.5; DB 2; Length 342;  
Best Local Similarity 68.8%; Pred. No. 1.5e-69;  
Matches 243; Conservative 32; Mismatches 65; Indels 13; Gaps 3;  
QY 2 SPTSPKVPFLSLCSTOPDGNVVIACLVQGFQEPPLSVTWSESGQGVTAARNPPSQDASG 61  
Db 2 SETSPKIFPLTLGSSSPAGVVIACLVROFFPPEPLTWSPREGVIVRNPPAQ--AG 59  
QY 62 DLYTSSQLTLPATOCCLACKSVTCHVKYTNPSODVTVPCVPSTPPTPSPPTSPS 121  
Db 60 GYLTWSSQTLTPEQCPAQDILKQVHLKSKSQSVNVPCKV-----LPSPCPQ 109  
QY 122 CCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFWTWPSGKSAVQGPDPDRDLG 181  
Db 110 CCKPSLSLOPPALADLLLGSNASLTCTLSGLKKSEGVSTWQPSGGKDAVQASPRDSCG 169  
QY 182 CYSVSVSLGSCAEPNHNKGTFTCTAAYPSKPTLTATLSK-SGNTFRPEVHLPPPSBEL 240  
Db 170 CYSVSSILPCADPMWKGSTFCTAAHSELKSGALTATITPKVNTFRPQVHLLPPPSBEL 229  
QY 241 ALNELVTLTCLARGSPKDVLRWLOGSQELPREKYLTVASRQEPGQGTTFVAVTSILRV 300  
Db 230 ALNELVTLTCLRVGSPKDVLRWLOGGQELPRDKYLVWESLPEPGQAIPVAVTSILRV 289  
QY 301 AAEDWKKGDTFCMWGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 290 DAEDWKGDTFCMWGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEAGICY 342  
RESULT 7  
C22360  
Ig alpha-2 chain C region (allotype A2m(2)) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 20-Jun-2000  
C:Accession: C22360  
R:Planagan, J.G.; Lefranc, M.P.; Rabbitts, T.H.  
Cell 36, 681-688, 1984  
A:Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1 an  
A:Reference number: A94653; MUID:84130179; PMID:6421489  
A:Accession: C22360  
A:Molecule type: DNA  
A:Residues: 1-220 <PLA>  
A:Cross-references: GB:AJ012264; NID:g3819787; PIDN:CAA09968.1; PID:g3819788  
C:Comment: This sequence revises entry A2HU (A02172).  
C:Genetics:  
A:Gene: GDB:IGHA2  
A:Cross-references: GDB:119333; OMIM:147000  
A:Map position: 14q32.33-14q32.33  
A:Introns: 90/1  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:110-182/Domain: immunoglobulin homology <IMM>  
Query Match 58.6%; Score 1113; DB 2; Length 220;  
Best Local Similarity 93.2%; Pred. No. 5.2e-60;  
Matches 205; Conservative 10; Mismatches 5; Indels 0; Gaps 0;  
QY 134 LEDLLGSEANLTCTLTGLRDASGVTFWTWPSGKSAVQGPDPDRDLGCGYSVSVSLGCA 193  
Db 1 LEDLLGSEANLTCTLTGLRDASGATFTWTPSSGKSAVQGPDPDRDLGCGYSVSVLPFGCA 60  
QY 194 EPWNHCKTFTCTAAYPSKPTLTATLSKSGNTFRPEVHLPPPSBELALNELVTLTCLAR 253  
Db 61 QPNHGETTCTTAHPELKTPLTANITKSGNTFRPEVHLPPPSBELALNELVTLTCLAR 120  
QY 254 GFSPKDVLRWLOGSQELPREKYLTVASRQEPGQGTTFVAVTSILRVAAEDWKKGDTFSC 313  
C

Db 121 GFSPKDVLRWLOGSQELPREKYLTVASRQEPGQGTTFVAVTSILRVAAEDWKKGDTFSC 180  
QY 314 MVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 181 MVGHEALPLAFTOKTIDRMAGKPTHINVSVMMAEADGTCY 220  
RESULT 8  
A45966  
Ig alpha chain C region - shrew mouse  
C:Species: Mus pahari  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 16-Jul-1999  
C:Accession: A45966  
R:Osborne, B.A.; Golde, T.E.; Schwartz, R.L.; Rudikoff, S.  
Genetics 119, 925-931, 1988  
A:Title: Evolution of the Iga heavy chain gene in the genus Mus.  
A:Reference number: A45966; MUID:88313645; PMID:2842228  
A:Accession: A45966  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-342 <OSB>  
A:Cross-references: GB:I13020  
A:Note: the authors translated the codon AAC for residue 46 as Lys, ATG for residue 334 &  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:232-304/Domain: immunoglobulin homology <IMM>  
Query Match 58.6%; Score 1112.5; DB 2; Length 342;  
Best Local Similarity 58.7%; Pred. No. 8.6e-60;  
Matches 209; Conservative 57; Mismatches 71; Indels 19; Gaps 5;  
QY 2 SPTSPKVPFLSL---CSTOPDGNVVIACLVQGFQEPPLSVTWSESGQGVTAARNPPSQD 58  
Db 2 SPRNPITYPLTLPPALSSEP---VIIGCLIHDFPSGTMNVTWGRSGNDITTVNFPAL- 57  
QY 59 ASGDLYTSSQLTLPATOCCLACKSVTCHVKYTNPSQDVTVPCVPSTPPTSPSTPTPT 118  
Db 58 ASGGRYTWSSQLTLPKECPAGESVKCSVQHDSPNPQVLELVKC-----SEPPLP 106  
QY 119 SPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTFWTWPSGKSAVQGPDPDRD 178  
Db 107 PSTICQPSLSLQRPALDILLGSDASLTCTLSGLKSTEGWFTWETTTGKDAVQKKPVQD 166  
QY 179 LGCYSVSVSLGSCAEPNHNKGTFTCTAAYPSKPTLTATLSK-SGNTFRPEVHLPPPS 237  
Db 167 SCGCFSSVSVLPFCARWNSGASFTCTVTHPDSPTLTGTIAKVTVNTFPQVHLVPPPS 226  
QY 238 BELALNELVTLTCLARGSPKDVLRWLOGSQELPREKYLTVASRQEPGQGTTFVAVTSI 297  
Db 227 BELALNELVSLTCLVRAFNPKEVLRWLHGNEELSPESYLVPFELKEPGEGATTVLTVSV 286  
QY 298 LRVAEDWKKGDTFCMWGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 287 LRVSATWKGQDQYSCMWGHEALPMNFTOKTIDRLSGKPTNVSVSVIMSEGDCICY 342  
RESULT 9  
AHMS  
Ig alpha chain C region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Dec-1981 #sequence\_revision 18-Dec-1981 #text\_change 09-Jul-2004  
C:Accession: A91479; A92245; A93857; A02173  
R:Auffray, C.; Nagotte, R.; Sikorav, J.L.; Heidmann, O.; Rougeon, F.  
Gene 13, 365-374, 1981  
A:Title: Mouse immunoglobulin A: nucleotide sequence of the structural gene for the alpha  
A:Reference number: A91479; MUID:81261947; PMID:6790349  
A:Contents: myelomas ABE48 and J558  
A:Accession: A91479  
A:Molecule type: mRNA  
A:Residues: 1-344 <AUF>  
A:Cross-references: UNIPROT:P01878  
R:Robinson, E.A.; Appella, E.  
J. Biol. Chem. 254, 11418-11430, 1979

A;Title: Amino acid sequence of a mouse myeloma immunoglobulin heavy chain (MOPC 47A) with  
A;Reference number: A92245; MUID:80049769; PMID:115869  
A;Contents: MOPC 47A  
A;Accession: A92245  
A;Molecule type: protein  
A;Residues: 1,'A',3-17,'C',19-66,'S',68-72,'T',74-134,'Q',136-140,'D',142-167,'E',169-211  
A;Note: the final C-region domain is deleted from this chain  
A;Note: Cys-18 may participate in the heavy-light chain bond  
R;Robinson, E.A.; Appella, E.  
Proc. Natl. Acad. Sci. U.S.A. 77, 4909-4913, 1980  
A;Title: Complete amino acid sequence of a mouse immunoglobulin alpha chain (MOPC 511).  
A;Reference number: A93857; MUID:81054880; PMID:6776528  
A;Contents: M511  
A;Accession: A93857  
A;Molecule type: protein  
A;Residues: 1-17,'C',19-66,'S',68-111,'G',113-134,'Q',136-140,'D',142-234,'G',236-254;29  
A;Note: this chain appears to lack residues 255-290  
A;Note: the sequence is compared with that of mouse MOPC 47A, and a genetic mechanism fo  
A;Note: this chain was isolated from a myeloma protein that binds phosphorylcholine  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F;131-197/Domain: immunoglobulin homology <IM1>  
F;234-306/Domain: immunoglobulin homology <IM2>  
F;26-84,76-100,114-171,138-195/Disulfide bonds: #status predicted  
F;38,329/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F;101/Binding site: carbohydrate (Ser) (covalent) #status experimental  
  
Query Match 56.2%; Score 1066.5; DB 1; Length 344;  
Best Local Similarity 58.4%; Pred. No. 4.9e-57;  
Matches 206; Conservative 53; Mismatches 83; Indels 11; Gaps 6;  
  
QY 2 SPTSPKVFPLSLCSTQPDGNVIA-CLVQGFPPQEPPLSVTWSGSGVGTARNFPSPQDASG 61  
DB 2 SARNPTIPLTPALSDPVIICLIHDYFPSTMTWTKSGKQDITVNFPPAL-ASG 60  
  
QY 62 DLYTTSSQLTLPATQCLACKSVTCHVKYTNPSQDVTPCPVSTPTPTSPSTPTSPS 121  
DB 61 GRVTMSNQLTLPVCEPESGKSVQSDGNSFVQELDVNCSGP-TTP-----PPITIPS 113  
  
QY 122 CCHPRLSLHRLPALEDDLGLSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLG 181  
DB 114 -CQPSLSLQRPALDDLGLSDASITCTLGLRNPEGAVFTWEPSTGDAVQKAVQNSCG 172  
  
QY 182 CYSVSSVSLGCAEPNHNKGTFTCTAAPE-SKTPLTATLSK-SGNTFRPEVHLPPSEEL 240  
DB 173 CYSVSSVLPGCAERNKSGAFKCTVTHPESGT-LTGIATKVTNTVFPQVHLPPSEEL 231  
  
QY 241 ALNELVTLTCLARGSPKDVLRWLOGSQELPREKYLTVASRQPSQGTTFFAVTSILRV 300  
DB 232 ALNELLSLTCLVRAFNPKEVLVRLHNGNEELSPESYLVPFELKEPGEATVLTSLRV 291  
  
QY 301 AEDWKGDFTSCWVGHEALPLAFTQKTIDRLACKPHTVNVSVVMAEVDGTCY 353  
DB 292 SAETWKQGDQYSCWVGHEALPMNFTQKTIDRLSGKPTNVSVVIMSGDGTIC 344  
  
RESULT 10  
S09267  
Ig alpha chain C region - rabbit (fragment)  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 16-Jul-1999  
C;Accession: S09267  
R;Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.  
EMBO J. 8, 4041-4047, 1989  
A;Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13  
A;Reference number: S09264; MUID:90076124; PMID:2512120  
A;Accession: S09267  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-357 <BUR>  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: immunoglobulin  
F;142-208/Domain: immunoglobulin homology <IMM>  
  
Query Match 53.8%; Score 1022; DB 2; Length 357;  
Best Local Similarity 56.9%; Pred. No. 2.4e-54;  
Matches 203; Conservative 35; Mismatches 109; Indels 10; Gaps 4;  
  
QY 4 TSPKVFPLSLCSTQPDGNVIA-CLVQGFPPQEPPLSVTWSGSGVGTARNFPSPQDASGL 63  
DB 4 TRPILPLSPILPGEPVVGILRGFPPLPLSVTWTNSGENLT---FPPVQSATSSL 60  
  
QY 64 YTTSSQLTLPATQCLACKSVTCHVKYTNPSQDVTPCPVSTPTPTSPSTP-----PTP 118  
DB 61 YTTCSLLRLLAECPEENSVACHVEHNYDKGVTVTPSPPECOPTTPGSDTTCPCPCP 120  
  
QY 119 SPSCCHPRLSLHRLPALEDDLGLSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRD 178  
DB 121 SPSCGFPPLSLQRPFLRDLNLSNASLTCTLGLKNPEGAVFTWEPNGKPKVQOSVQSY 180  
  
QY 179 LCGCYSVSSVSLGCAEPNHNKGTFTCTAAPE-SKTPLTATLSK-SGNTFRPEVHLPPP 236  
DB 181 PCGCYSVSSVLPGCAEPNAGTEFTCTVTHPEIEGGPLTAKISKDTGAILPPQVHLPPP 240  
  
QY 237 SEELALNELVTLTCLARGSPKDVLRWLOGSQELPREKYLTVASRQPSQGTTFFAVTS 296  
DB 241 SEELALNELVTLTCLVRGSPKDVLRVYTNKGVNVPENSLVWPKLPPEPQOEPTTYAVTS 300

C;Keywords: immunoglobulin  
F;142-208/Domain: immunoglobulin homology <IMM>

Query Match 54.6%; Score 1035.5; DB 2; Length 357;  
Best Local Similarity 57.1%; Pred. No. 3.7e-55;  
Matches 205; Conservative 43; Mismatches 98; Indels 13; Gaps 8;

QY 4 TSPKVFPLSL--CSTQPDGNVIA-CLVQGFPPQEPPLSVTWSGSGVGTARNFPSPQDAS 60  
DB 3 TPPIRPLTCPGCVLKDTSATIVAGCLIRGFFPRGPLGVTWMDNRANLT---FPPVQSAT 59  
  
QY 61 GDYITSSQLTLPATQCLACKSVTCHVKYTNPSQDVTPCPV---PSTPTPSPTP-P 116  
DB 60 SSLYTTCVSLSPAEQCPAGNSVACRVEH--NNKQDLTVPCLACNKTIEPTPTKPCP 118  
  
QY 117 TPSPSCCHPRLSLHRLPALEDDLGLSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDP 176  
DB 119 CPSPSCOKPRLSLQRPDLGLDLSNASLTCTLGLLNPEGAVFTWNTWNGKEFVQSSAQ 178  
  
QY 177 RDLGCVSVSSVSLGCAEPNHNKGTFTCTAAPE-SKTPLTATLSK-SGNTFRPEVHLPP 234  
DB 179 RDHCGCVSVSSVLPGCAEPNAGTVFTCTVTHPEIDSGSLTATISKDTGSLIPQVHLPP 238  
  
QY 235 PPSEELALNELVTLTCLARGSPKDVLRWLOGSQELPREKYLTVASRQPSQGTTFFAV 294  
DB 239 PPSEELALNALVTLTCLVRGSPKDVLRVYTNKGLQVPKDSFLVWPKLPPEPQOEPTTYAV 298  
  
QY 295 TSLRVAEEDWKGDFTSCWVGHEALPLAFTQKTIDRLACKPHTVNVSVVMAEVDGTCY 353  
DB 299 TSLRVPADENWQNESYTCVVGHEGLAEHFTQKTIDRLACKPHTVNVSVVADVEGVCY 357

## RESULT 11

S09269

Ig alpha chain C region - rabbit (fragment)

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 16-Jul-1999

C;Accession: S09269

R;Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.

EMBO J. 8, 4041-4047, 1989

A;Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13

A;Reference number: S09264; MUID:90076124; PMID:2512120

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-357 &lt;BUR&gt;

C;Superfamily: immunoglobulin C region; immunoglobulin homology

C;Keywords: immunoglobulin

F;142-208/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 53.8%; Score 1022; DB 2; Length 357;  
Best Local Similarity 56.9%; Pred. No. 2.4e-54;  
Matches 203; Conservative 35; Mismatches 109; Indels 10; Gaps 4;

QY 4 TSPKVFPLSLCSTQPDGNVIA-CLVQGFPPQEPPLSVTWSGSGVGTARNFPSPQDASGL 63  
DB 4 TRPILPLSPILPGEPVVGILRGFPPLPLSVTWTNSGENLT---FPPVQSATSSL 60  
  
QY 64 YTTSSQLTLPATQCLACKSVTCHVKYTNPSQDVTPCPVSTPTPTSPSTP-----PTP 118  
DB 61 YTTCSLLRLLAECPEENSVACHVEHNYDKGVTVTPSPPECOPTTPGSDTTCPCPCP 120  
  
QY 119 SPSCCHPRLSLHRLPALEDDLGLSEANLTCTLTGLRDASGVTFTWTPSSGKSAVQGPDRD 178  
DB 121 SPSCGFPPLSLQRPFLRDLNLSNASLTCTLGLKNPEGAVFTWEPNGKPKVQOSVQSY 180  
  
QY 179 LCGCYSVSSVSLGCAEPNHNKGTFTCTAAPE-SKTPLTATLSK-SGNTFRPEVHLPPP 236  
DB 181 PCGCYSVSSVLPGCAEPNAGTEFTCTVTHPEIEGGPLTAKISKDTGAILPPQVHLPPP 240  
  
QY 237 SEELALNELVTLTCLARGSPKDVLRWLOGSQELPREKYLTVASRQPSQGTTFFAVTS 296  
DB 241 SEELALNELVTLTCLVRGSPKDVLRVYTNKGVNVPENSLVWPKLPPEPQOEPTTYAVTS 300

```
QY 297 ILRVAEDWKGDFTSCVMGHEALPLAFTQKTIDRLAGKPTHVNVSVMAEVDGTCY 353
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 LURVPAEDWNQNESYTCVVGHEGLAEHFTQRTIDRLSGRPTHVNVSVVADVEGVY 357

RESULT 12
S09265
Ig alpha chain C region - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09265
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference number: S09264; MUID:90076124; PMID:2512120
A:Accession: S09265
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-357 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:246-319/Domain: immunoglobulin homology <IMM>

Query Match 53.5%; Score 1016; DB 2; Length 357;
Best Local Similarity 57.3%; Pred. No. 5 4e-54;
Matches 205; Conservative 35; Mismatches 106; Indels 12; Gaps 6;

QY 4 TSPKVPFLSLCSTQPDGNVVIACLVGQFPQBPPLSVTWSESGQGVTAARNFPPSQDASGDL 63
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 SSPDLPLPCPILEPCEPMVIGCLIRGFFPRGPLVTVMNVGESVI---FPPVPSPSSL 60

QY 64 YTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCPVPSTPTPTSPSTPT-----P 118
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 YTTSLRLPAEQCPENSAQCRVEH-NNKGDQVTPSPACNVESTIEPTPTTCPCPCP 119

QY 119 SPSCCHPRLSLRPALEDLLLGSEANLTCTLTGLRDASGVTFWTTPSSGKSAVQGPDRD 178
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 SPSCGKPSLSLQRPDLGDLJLNSASLTCTLTGLLNPEGAFTWPTTFKEPVOQSPQLD 179

QY 179 LCCGYSVSVSLGCAEPMNHGKTFTCTAAYPESK-TPLTATLSK-SGNTFRPEVHLPPP 236
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 HCGCYSSVSLPGCAVLNAGTEFTCTVTHPETEGSLGTISKDGTSLIPQVHLLPPP 239

QY 237 SEELALNELVTLTCLARGSPKDVLRWL-QSQEPLPREKYLTVASRQPSQGTTFATV 295
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 SEELALNLTTLCLVRGSPKDVLSWTHNGTPVVPKDSYLWKPLRBPFGQDPTTYAIT 299

QY 296 SILRVAEDWKGDFTSCVMGHEALPLAFTQKTIDRLAGKPTHVNVSVMAEVDGTCY 353
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 SLURVPAEDWNQDYSYCVVGHEGLAEHFTQKTIDRLAGKPTHVNVSVVADVEGVY 357

RESULT 13
S09268
Ig alpha chain C region - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09268
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference number: S09264; MUID:90076124; PMID:2512120
A:Accession: S09268
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-358 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:248-320/Domain: immunoglobulin homology <IMM>

Query Match 51.9%; Score 986; DB 2; Length 358;
Best Local Similarity 55.4%; Pred. No. 3.4e-52;
```

```
Matches 199; Conservative 41; Mismatches 105; Indels 14; Gaps 7;

QY 5 SPKVPFL--SLCSTQPDGNVITA-CLVQGFPPQBPPLSVTWSESGQGVTAARNFPPSQDASG 61
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 SPRLPFLIHPRCALKDDTSATVIAGCLIRGFFPLGLPFLSVSNASGKNVT---FPPVPSTGS 60

QY 62 DLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCPVPSTPTPTSPSTPT-----P 116
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GPYTTCSSLTPEQCPEDDNNVCHVEHNYDKQNLTVLP-ECOPFTPTPTTCPCP 119

QY 117 TPSPCCCHPRLSLRPALEDLLLGSEANLTCTLTGLRDASGVTFWTTPSSGKSAVQGPDP 176
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 CPSPSCGEPSSLQRPDLGDLJLNSASLTCTLTGLLDPEGAFTWPTTFKEPVOQLSPK 179

QY 177 RLCCGYSVSVSLGCAEPMNHGKTFTCTAAYPESK-TPLTATLSK-SGNTFRPEVHLPL 234
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 LDHCGCYSSVSLPGCAAAWNAAGTKFNCTVTHPEIKGVSLTDIIISKDTGVVIAPOVHLLP 239

QY 235 PSEELALNELVTLTCLARGSPKDVLRWLQGSQBPPLPREKYLTVASRQPSQGTTFPAV 294
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 PPSDELALNALVTLTCLVRGSPKDVLYVWTKNGVEVPKDSFLVWKPLPEPQGEPTTYAV 299

QY 295 SILRVAEDWKGDFTSCVMGHEALPLAFTQKTIDRLAGKPTHVNVSVMAEVDGTCY 353
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 TSLLRVPAEDWNQNESYTCVVGHEGLAEHFTQRTIDRLAGRPTHVNVSVVADVEGVY 358

RESULT 14
S09274
Ig alpha chain C region - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09274
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference number: S09264; MUID:90076124; PMID:2512120
A:Accession: S09274
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-347 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:132-198/Domain: immunoglobulin homology <IMM>

Query Match 51.8%; Score 984; DB 2; Length 347;
Best Local Similarity 54.4%; Pred. No. 4.3e-52;
Matches 193; Conservative 45; Mismatches 103; Indels 14; Gaps 4;

QY 4 TSPKVPFLSLCSTQPDGN---VVIACLVGQFPQBPPLSVTWSESGQGVTAARNFPPSQDAS 60
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 TTPGIYPLPLRVSDGNSQTVVVGCLIRGFFPLGLPLRVSMNVSRNVSIIYFPPTPTGT 61

QY 61 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTPCPVPSTPTPTSPSTPTTPSP 120
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 SGFYTACSELILPDTCLEYDSAAACHVEYNVSNESLVPFPDP-----CEQCHCP 112

QY 121 SCCHPRLSLRPALEDLLLGSEANLTCTLTGLRDASGVTFWTTPSSGKSAVQGPDRDL 180
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 113 SCEEPSLSLQRPDLRLDLGSDASLTCTLTURGLKDEGAFTWGTNGNEPVOQSPQDPC 172

QY 181 GCYSVSVSLGCAEPMNHGKTFTCTAAYPESK-TPLTATLSK-SGNTFRPEVHLPPPSE 238
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 173 GCYSVSVSLPGCAEPMNHGKTFTCTVTHPEIGSSLTATISKDTGSLTLPVHLLPPPSE 232

QY 239 ELALNELVTLTCLARGSPKDVLRWLQGSQBPPLPREKYLTVASRQPSQGTTFATVSTIL 298
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 233 ELALNALVTLTCLVRGSPKDVLSWTKGVKVPENSFLVWKPLPEPQGDPTTYAVTSL 292

QY 299 RVAEDWKGDFTSCVMGHEALPLAFTQKTIDRLAGKPTHVNVSVMAEVDGTCY 353
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 293 RVPADWNQNESYSCVVAHEGLAEHFTQRTIDRLAGRPTHVNVSVVADVEAVCY 347
```

RESULT 15  
S09270  
Ig alpha chain C region - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 16-Jul-1999  
C:Accession: S09270  
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.  
EMBO J. 8, 4041-4047, 1989  
A:Title: The IgA heavy-chain gene family in rabbit: cloning and sequence analysis of 13  
A:Reference number: S09264; MURD:90076124; PMID:2512120  
A:Accession: S09270  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-348 <BUR>  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:133-199/Domain: immunoglobulin homology <IMM>

Query Match 51.6%; Score 980; DB 2; Length 348;  
Best Local Similarity 54.0%; Pred. No. 7.5e-52;  
Matches 191; Conservative 47; Mismatches 102; Indels 14; Gaps 4;

Qy	5	SPKVFPLSLCSTQPDGN--VVIACLVQGGPPQEPPLSVTWSESGQGVTAENFPESQDASG	61
Db	4	TPDIFPLNLPVRVSDGNSQTVVVGCLIRGFFPPPLRVSNVSNRENMSVNFPPAPTGT	63
Qy	62	DLVTTSQLTLPATQCLAGKSVTCHVRKHTNPNSQDVTVPQVPSPTPTSPSTPTSPS	121
Db	64	GPYTACSELILPVTOCLEYDSAAACHVEYNSVINESLVPFPDP-----CEQCHCPS	114
Qy	122	CCHPRLSLHRLPALEDLLGSEANLTCTLTGLRDASGVFTWTPSSGKSAVQGGPDRDL	181
Db	115	CEEPSLSLQRPDLRLDLLGSDASLTCTLRGLKYPEDAVFTWEPTNGNEFVQSPQRP	174
Qy	182	CYSVSSVLSGCAEPWNHKGTFCTCTAAYPESK-TPLTATLSK-SCNTERPEVHLLPP	239
Db	175	CYSVSSVLPGCAEPWNAGTEFTCTVTHPEIEGGSLTATISKOTGSLTPPQVHLLPP	234
Qy	240	LALNELVTLTCLARGFSPKDVLRWLGSOELPREKYLTVASROEPSQGTTFITAVTS	299
Db	235	LALNALVTLTCLVRGFSFKDVLVYWRKDVPEPNSFLVWKPLPEFGQDPTTYAVTS	294
Qy	300	VAAEDMKKGDTFCMVGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTCY	353
Db	295	VSAEDWNQGDSTYTCVVGHEGLAEHFTQRTIDREAGRPTHVNVSVVWADVEAVCY	348

Search completed: October 25, 2005, 11:08:26  
Job time : 13.158 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 10:56:47 ; Search time 50.728 Seconds  
(without alignments)  
3563.394 Million cell updates/sec

Title: US-10-644-256-3\_COPY\_137\_489

Perfect score: 1898

Sequence: 1 ASPTSPKVPFLSLCSTQPDG.....GKPTHVNVSVMAEVDGTCY 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1894	99.8	500	2 Q9BRV0	Q9brv0 homo sapien
2	1893	99.7	494	2 Q6ZWE4	Q6zwe4 homo sapien
3	1893	99.7	496	2 Q96DK0	Q96dk0 homo sapien
4	1893	99.7	499	2 Q8N5K4	Q8n5k4 homo sapien
5	1889	99.5	353	1 ALC1 HUMAN	P01876 homo sapien
6	1889	99.5	384	2 Q9UPF0	Q9upf0 homo sapien
7	1889	99.5	493	2 Q8NCL6	Q8nc16 homo sapien
8	1889	99.5	493	2 Q6GMX2	Q6gmx2 homo sapien
9	1889	99.5	494	2 Q96K68	Q96k68 homo sapien
10	1889	99.5	496	2 Q96KX8	Q96kx8 homo sapien
11	1889	99.5	506	2 Q6MZW0	Q6mzw0 homo sapien
12	1889	99.5	519	2 Q6N092	Q6n092 homo sapien
13	1884	99.3	506	2 Q6N090	Q6n090 homo sapien
14	1881	99.1	497	2 Q8WY24	Q8wy24 homo sapien
15	1842	97.0	353	1 ALC1 GORGO	P20758 gorilla gor
16	1710.5	90.1	340	1 ALC2 HUMAN	P01877 homo sapien
17	1710.5	90.1	416	2 Q9NPF6	Q9npp6 homo sapien
18	1710.5	90.1	477	2 Q6GMX7	Q6gmx7 homo sapien
19	1710.5	90.1	478	2 Q6NYH3	Q6nyh3 homo sapien
20	1710.5	90.1	479	2 Q6MZW6	Q6mzw6 homo sapien
21	1710.5	90.1	480	2 Q6P089	Q6p089 homo sapien
22	1710.5	90.1	487	2 Q6ZVX0	Q6zvx0 homo sapien
23	1704.5	89.8	478	2 Q7Z379	Q7z379 homo sapien
24	1704.5	89.8	492	2 Q7Z374	Q7z374 homo sapien
25	1701.5	89.6	483	2 Q6MZX9	Q6mzx9 homo sapien
26	1667.5	87.9	498	2 Q6N041	Q6n041 homo sapien
27	1660.5	87.5	500	2 Q6N091	Q6n091 homo sapien
28	1088.5	57.3	488	2 Q91WR1	Q91wr1 mus musculus
29	1088.5	57.3	489	2 Q8VCX4	Q8vcx4 mus musculus
30	1088	57.3	481	2 Q91WT1	Q91wt1 mus musculus
31	1088	57.3	481	2 Q91WT3	Q91wt3 mus musculus

32	1088	57.3	482	2 Q91X92	Q91x92 mus musculus
33	1088	57.3	482	2 Q8K172	Q8k172 mus musculus
34	1088	57.3	484	2 Q8VEA0	Q8vea0 mus musculus
35	1088	57.3	485	2 Q6PDB8	Q6pdb8 mus musculus
36	1088	57.3	487	2 Q80Z17	Q80zi7 mus musculus
37	1088	57.3	488	2 Q8KQF2	Q8kof2 mus musculus
38	1086	57.2	479	2 Q91WP5	Q91wp5 mus musculus
39	1085	57.2	480	2 Q91XE1	Q91xe1 mus musculus
40	1082.5	57.0	486	2 Q91Z07	Q91z07 mus musculus
41	1082.5	57.0	487	2 Q99KA4	Q99ka4 mus musculus
42	1079.5	56.9	426	2 Q9DCD9	Q9dcd9 mus musculus
43	1069.5	56.3	479	2 Q99M22	Q99m22 mus musculus
44	1069.5	56.3	479	2 Q7TWK4	Q7tmk4 mus musculus
45	1069.5	56.3	480	2 Q8K0Z4	Q8k0z4 mus musculus

#### ALIGNMENTS

#### RESULT 1

ID	Q9BRV0	PRELIMINARY;	PRT;	500 AA.
AC	Q9BRV0;			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE	MGC27165 protein.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RS	TISSUE=Prostate;			
RC	MDLLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RN	[2]			
RP	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RS	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RA	Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC005951; AAH05951.1; -;			
DR	HSSP; P01876; IOWO.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003597; Ig cl.			
DR	InterPro; IPR003006; Ig_WHC.			
DR	InterPro; IPR003596; Ig_v.			
DR	Pfam; PF07654; Cl-set; 2.			
DR	Pfam; PF00047; Ig; 1.			
DR	SMART; SM00406; IGV; 1.			
DR	PROSITE; PS0835; IG_LIKE; 4.			
DR	PROSITE; PS00290; IG_WHC; UNKNOWN 1.			
DR	SEQUENCE 500 AA; 54154 MW; 0A9BF43F2A3CC6D9 CRC64;			

```
Query Match          99.8%; Score 1894; DB 2; Length 500;
Best Local Similarity 99.7%; Pred. No. 7.5e-112;
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPPPSQDAS 60
    |||
Db 148 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPPPSQDAS 207
    |||

QY 61 GDLYTTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
    |||
Db 208 GDLYTTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 267
    |||

QY 121 SCCHPRLSLHRRPALEDLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 180
    |||
Db 268 SCCHPRLSLHRRPALEDLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 327
    |||

QY 181 GCVSVSVSLGCAEPNNHGKTFCTTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240
    |||
Db 328 GCVSVSVSLGCAEPNNHGKTFCTTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 387
    |||

QY 241 ALNELVTLTCLARGSPKDVLRWLGSOELPREKYLTVASRQEPSQGTTFVAVTSILRV 300
    |||
Db 388 ALNELVTLTCLARGSPKDVLRWLGSOELPREKYLTVASRQEPSQGTTFVAVTSILRV 447
    |||

QY 301 AAEDMKKGDTFSCMGVGEALPLAFTQKTIIDRLAGKPTHVNVSVMAEVDGTCY 353
    |||
Db 448 AAEDMKKGDTFSCMGVGEALPLAFTQKTIIDRLAGKPTHVNVSVMAEVDGTCY 500
    |||

RESULT 2
Q6ZW64 PRELIMINARY; PRT; 494 AA.
AC Q6ZW64
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Hypothetical protein FLJ41552.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
RA Irie R., Oeuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saiko K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuna M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuho Y., Nagai K., Isogai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK123546; BAC85641.1; -.
DR HSSP; P01842; 1AOK.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG-cl.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF07654; Cl-set; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGC1; 2.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
SQ SEQUENCE 494 AA; 53321 MW; CD8140B37A23A882 CRC64;

Query Match          99.7%; Score 1893; DB 2; Length 494;
Best Local Similarity 99.7%; Pred. No. 8.6e-112;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPPPSQDAS 60
    |||
Db 144 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPPPSQDAS 203
    |||

QY 61 GDLYTTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
    |||
Db 204 GDLYTTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 263
    |||

QY 121 SCCHPRLSLHRRPALEDLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 180
    |||
```

```
Db 142 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPPPSQDAS 201
    |||
QY 61 GDLYTTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
    |||
Db 202 GDLYTTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 261
    |||

QY 121 SCCHPRLSLHRRPALEDLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 180
    |||
Db 262 SCCHPRLSLHRRPALEDLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 321
    |||

QY 181 GCVSVSVSLGCAEPNNHGKTFCTTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 240
    |||
Db 322 GCVSVSVSLGCAEPNNHGKTFCTTAAYPESKTPLTATLSKSGNTFRPEVHLLPPSEEL 381
    |||

QY 241 ALNELVTLTCLARGSPKDVLRWLGSOELPREKYLTVASRQEPSQGTTFVAVTSILRV 300
    |||
Db 382 ALNELVTLTCLARGSPKDVLRWLGSOELPREKYLTVASRQEPSQGTTFVAVTSILRV 441
    |||

QY 301 AAEDMKKGDTFSCMGVGEALPLAFTQKTIIDRLAGKPTHVNVSVMAEVDGTCY 353
    |||
Db 442 AAEDMKKGDTFSCMGVGEALPLAFTQKTIIDRLAGKPTHVNVSVMAEVDGTCY 494
    |||

RESULT 3
Q96DK0 PRELIMINARY; PRT; 496 AA.
AC Q96DK0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein FLJ25298.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach mucosa;
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Kawamura M.,
RA Suzuki Y., Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA Kawakami B., Nagai K., Isogai T., Sugano S.,
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK058027; BAB71633.1; -.
DR HSSP; P01876; 1OWO.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG-cl.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF07654; Cl-set; 2.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
SQ SEQUENCE 496 AA; 53532 MW; C72EE1E247C86FED CRC64;

Query Match          99.7%; Score 1893; DB 2; Length 496;
Best Local Similarity 99.7%; Pred. No. 8.6e-112;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPPPSQDAS 60
    |||
Db 144 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGOGVTARNPPPSQDAS 203
    |||

QY 61 GDLYTTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 120
    |||
Db 204 GDLYTTSSQLTLPATOCCLAGKSVTCHVKHYTNPSQDVTVPVPSPTPTPTSP 263
    |||

QY 121 SCCHPRLSLHRRPALEDLLGSEANLTCTLTGLRDASGVFTTWPSSGKSAVOGPPDRDLC 180
    |||
```

Db 264 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDL 323  
 QY 181 GCYSVSVSLVSGCAEPNHNHGTFTCTAAYPSKTPLTATLSKSGNTRPRPEVHLLPPSEEL 240  
 Db 324 GCYSVSVSLVSGCAEPNHNHGTFTCTAAYPSKTPLTATLSKSGNTRPRPEVHLLPPSEEL 383  
 QY 241 ALNELVTLTCLARGFSPKDVLRWLOGSQBELPREKYLTVASROPSQGTTFPAVTSILRV 300  
 Db 384 ALNELVTLTCLARGFSPKDVLRWLOGSQBELPREKYLTVASROPSQGTTFPAVTSILRV 443  
 QY 301 AEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
 Db 444 AEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 496

RESULT 4  
 Q8NSK4 ID Q8NSK4 PRELIMINARY; PRT; 499 AA.  
 AC Q8NSK4; 2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE MGC27165 protein.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA Strausberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC032249; AAH32249.1; -;  
 DR HSSP; P01876; 10W0.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig-cl.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF07654; C1-set; 2.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN 1.  
 SQ SEQUENCE 499 AA; 93376 MW; 93A5C89582054F32 CRC64;

Query Match 99.78; Score 1893; DB 2; Length 499;  
 Best Local Similarity 99.78; Pred No. 8.7e-112;  
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTOPDGNVVIACLVQGFPOEPLSVTWSESGQGVTTARNPPSQDAS 60  
 Db 147 ASPTSPKVPFLSLCSTOPDGNVVIACLVQGFPOEPLSVTWSESGQGVTTARNPPSQDAS 206  
 QY 61 GDLYTTSSQLTLPATOCLAGKSVTCHVKHYTNPSQDVTVPCVPSTPTPTSPSTPTPTSP 120  
 Db 207 GDLYTTSSQLTLPATOCLAGKSVTCHVKHYTNPSQDVTVPCVPSTPTPTSPSTPTPTSP 266  
 QY 121 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDL 180  
 Db 267 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDL 326  
 QY 181 GCYSVSVSLVSGCAEPNHNHGTFTCTAAYPSKTPLTATLSKSGNTRPRPEVHLLPPSEEL 240  
 Db 327 GCYSVSVSLVSGCAEPNHNHGTFTCTAAYPSKTPLTATLSKSGNTRPRPEVHLLPPSEEL 386  
 QY 241 ALNELVTLTCLARGFSPKDVLRWLOGSQBELPREKYLTVASROPSQGTTFPAVTSILRV 300  
 Db 387 ALNELVTLTCLARGFSPKDVLRWLOGSQBELPREKYLTVASROPSQGTTFPAVTSILRV 446  
 QY 301 AEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
 Db 447 AEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 499

RESULT 5  
 ALC1\_HUMAN ID ALC1\_HUMAN STANDARD; PRT; 353 AA.  
 AC P01876;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 25-JAN-2005 (Rel. 46, Last annotation update)  
 DE Ig alpha-1 chain C region.  
 GN NamesIGHA1;  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84130179; PubMed=6421489; DOI=10.1016/0092-8674(84)90348-9;  
 RA Flanagan J.G., Lefranc M.-P., Rabbitts T.H.;  
 RT "Mechanisms of divergence and convergence of the human immunoglobulin  
 RL alpha 1 and alpha 2 constant region gene sequences."  
 RL Cell 36:681-688(1984).  
 RN [2]  
 RP SEQUENCE (MYELOMA PROTEIN BUR), AND DISULFIDE BONDS.  
 RX MEDLINE=79151016; PubMed=107164;  
 RA Putnam F.W., Liu Y.-S.V., Low T.L.K.;  
 RT "Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal  
 RL IgA1 protease, digestion, Fab and Fc fragments, and the complete amino  
 RL acid sequence of the alpha 1 heavy chain."  
 RL J. Biol. Chem. 254:2865-2874(1979).  
 RN [3]  
 RP SEQUENCE (MYELOMA PROTEIN TRO).  
 RX MEDLINE=76023781; PubMed=809331;  
 RA Kratzin H., Altevogt P., Ruban E., Kortt A., Starosciak K.,  
 RA Hilschmann N.;  
 RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.).  
 RL II. The amino acid sequence of the H-chain, alpha-type, subgroup III;  
 RL structure of the complete IgA-molecule."  
 RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).  
 RN [4]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=80114124; PubMed=393607;  
 RA Yang C.-Y., Kratzin H., Gotz H., Hilschmann N.;  
 RT "Rule of antibody structure. Primary structure of a human monoclonal  
 RL IgA-immunoglobulin (myeloma protein Tro). VII. Purification and  
 RL characterization of the disulfide bridges."  
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:1919-1940(1979).  
 RN [5]  
 RP REVIEW.  
 RX MEDLINE=91054387; PubMed=2241915;



```
Db 272 ALNELVTLTCLARGFSPKDVLRWLQSQBQLPREKYLWTWASROEPSQGTTFFAVTSILRV 331
QY 301 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVMAEVDGTCY 353
Db 332 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVMAEVDGTCY 384

RESULT 7
Q8NCL6 PRELIMINARY; PRT; 493 AA.
AC Q8NCL6
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FJ90170.
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RA EMBL: AK074651; BAC11114.1; -
RA HSP: P01876; IOW0.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig-cl.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF07654; Cl-set; 2.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PSS0835; IG_LIKE; 4.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 493 AA; 53224 MW; 12ECDE094777101 CRC64;

Query Match 99.5%; Score 1889; DB 2; Length 493;
Best Local Similarity 99.4%; Pred. No. 1.5e-111;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGGFFQEPPLSVTWSSGQGVTAARNFPPSQDAS 60
Db 141 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGGFFQEPPLSVTWSSGQGVTAARNFPPSQDAS 200

QY 61 GDLYTSSQLTPATQCLAGKSVTCHVKHYTNPSQDVTVPSPSTPTPTSP 120
Db 201 GDLYTSSQLTPATQCLAGKSVTCHVKHYTNPSQDVTVPSPSTPTPTSP 260

QY 121 SCCHPRLSLHRPALEDLLGSEANLTCTGLRDASGVTFTWTPSSCKSAVQGPDRDLC 180
Db 261 SCCHPRLSLHRPALEDLLGSEANLTCTGLRDASGVTFTWTPSSCKSAVQGPDRDLC 320

QY 181 GCYSVSVSLSGCAEPWNHGKTFCTTAAYPSKTPLTATLSKSGNTFRPEVHLLPPPEEL 240
Db 321 GCYSVSVSLPGCAEPWNHGKTFCTTAAYPSKTPLTATLSKSGNTFRPEVHLLPPPEEL 380

QY 241 ALNELVTLTCLARGFSPKDVLRWLQSQBQLPREKYLWTWASROEPSQGTTFFAVTSILRV 300
Db 381 ALNELVTLTCLARGFSPKDVLRWLQSQBQLPREKYLWTWASROEPSQGTTFFAVTSILRV 440

QY 301 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVMAEVDGTCY 353
Db 441 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVMAEVDGTCY 493

RESULT 8
Q6GMX2 PRELIMINARY; PRT; 493 AA.
ID Q6GMX2
AC Q6GMX2;
```

```
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Phney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Kzyvynski M.I., Skalska U., Smalusz D.E., Schnertch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC073771; AAH73771.1; -
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig-cl.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF07654; Cl-set; 2.
DR SMART: SM00407; IG; 3.
DR SMART: SM00409; IG; 4.
DR SMART: SM00407; IGV; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PSS0835; IG_LIKE; 4.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 52865 MW; 55B999305B286203 CRC64;

Query Match 99.5%; Score 1889; DB 2; Length 493;
Best Local Similarity 99.4%; Pred. No. 1.5e-111;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGGFFQEPPLSVTWSSGQGVTAARNFPPSQDAS 60
Db 141 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGGFFQEPPLSVTWSSGQGVTAARNFPPSQDAS 200

QY 61 GDLYTSSQLTPATQCLAGKSVTCHVKHYTNPSQDVTVPSPSTPTPTSP 120
Db 201 GDLYTSSQLTPATQCLAGKSVTCHVKHYTNPSQDVTVPSPSTPTPTSP 260

QY 121 SCCHPRLSLHRPALEDLLGSEANLTCTGLRDASGVTFTWTPSSCKSAVQGPDRDLC 180
Db 261 SCCHPRLSLHRPALEDLLGSEANLTCTGLRDASGVTFTWTPSSCKSAVQGPDRDLC 320

QY 181 GCYSVSVSLSGCAEPWNHGKTFCTTAAYPSKTPLTATLSKSGNTFRPEVHLLPPPEEL 240
Db 321 GCYSVSVSLPGCAEPWNHGKTFCTTAAYPSKTPLTATLSKSGNTFRPEVHLLPPPEEL 380
```

```

QY 241 ALNELVTLTCLARGSPKDVLRWLGSGQELPREKYLTLWASRQEPSQGTTFITFAVTSILRV 300
DB 381 ALNELVTLTCLARGSPKDVLRWLGSGQELPREKYLTLWASRQEPSQGTTFITFAVTSILRV 440
QY 301 AAEDWKKGDTFSCMGVGHGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 353
DB 441 AAEDWKKGDTFSCMGVGHGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 493

RESULT 9
Q96K68 ID Q96K68 PRELIMINARY; PRT; 494 AA.
AC Q96K68;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ14473.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda K., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku T., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs."
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK027379; BAB55072.1; -.
DR FIP; S21205; S21205.
DR HSP; P01876; IOWO.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig.MHC.
DR InterPro; IPR003596; Ig.v.
DR Pfam; PF07654; C1-set; 2.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
SQ SEQUENCE 494 AA; 53088 MW; 9A1D7ABESAE4C0E CRC64;

Query Match 99.5%; Score 1889; DB 2; Length 494;
Best Local Similarity 99.4%; Pred. No. 1.5e-111;
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ASPTSPKVPPLSLCSTQPDGNVVIACLVOGFRPQBPPLSVTWSESGGVTAARNPPSODAS 60
DB 142 ASPTSPKVPPLSLCSTQPDGNVVIACLVOGFRPQBPPLSVTWSESGGVTAARNPPSODAS 201
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTSPSTPTSPSP 120
DB 202 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTSPSTPTSPSP 261
QY 121 SCCHPRLSLHRPALEDDLGLGSEANLTCTLTGLURDASGVTFITWTPSSGKSAVGPPDRDLIC 180
DB 262 SCCHPRLSLHRPALEDDLGLGSEANLTCTLTGLURDASGVTFITWTPSSGKSAVGPPDRDLIC 321
QY 181 GCYSVSSVLSGCAEPWNHGKFTCTAAVPESKTPLTATLSKSGNTFRPVRHLLPPSEEL 240
DB 322 GCYSVSSVLSGCAEPWNHGKFTCTAAVPESKTPLTATLSKSGNTFRPVRHLLPPSEEL 381
QY 241 ALNELVTLTCLARGSPKDVLRWLGSGQELPREKYLTLWASRQEPSQGTTFITFAVTSILRV 300
DB 382 ALNELVTLTCLARGSPKDVLRWLGSGQELPREKYLTLWASRQEPSQGTTFITFAVTSILRV 441
QY 301 AAEDWKKGDTFSCMGVGHGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 353
DB 442 AAEDWKKGDTFSCMGVGHGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 494

RESULT 10
Q96KX8 ID Q96KX8 PRELIMINARY; PRT; 496 AA.
AC Q96KX8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MGC27165 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klautner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016369; AAH16369.1; -.
DR HSP; P01876; IOWO.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig.MHC.
DR InterPro; IPR003596; Ig.v.
DR Pfam; PF07654; C1-set; 2.

```

```
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG_LIKE; 4.  
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.  
SQ SEQUENCE 496 AA; 53391 MW; D346929849040D69 CRC64;  
  
Query Match 99.5%; Score 1889; DB 2; Length 496;  
Best Local Similarity 99.4%; Pred. No. 1.6e-111;  
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60  
Db 144 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 203  
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKYTNPSQDVTPCPVPSPPTPTPTPTPSP 120  
Db 204 GDLYTTSSQLTLPATQCLAGKSVTCHVKYTNPSQDVTPCPVPSPPTPTPTPTPSP 263  
QY 121 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDPDLC 180  
Db 264 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDPDLC 323  
QY 181 GCYSVSVSLGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 240  
Db 324 GCYSVSVSLGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 383  
QY 241 ALNELVTLTCLARGFSPKDVLRWLOGSQELPREKYLTVASRQEPSQGTTFFAVTSILRV 300  
Db 384 ALNELVTLTCLARGFSPKDVLRWLOGSQELPREKYLTVASRQEPSQGTTFFAVTSILRV 443  
QY 301 AEDWKKGDFTFCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 444 AEDWKKGDFTFCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 496  
  
RESULT 11  
Q6MZWO PRELIMINARY; PRT; 506 AA.  
ID Q6MZWO  
AC Q6MZWO; 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein DKFZp686J11235 (Fragment).  
GN Name=DKFZp686J11235;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Human small intestine;  
RG The German Human cDNA Consortium;  
RA Bloeker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,  
RA Fobo G., Han M., Wiemann S.;  
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
EMBL; BX640847; CAB45917.1; -  
DR HSP; P01751; IAGW.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003597; IG-cl.  
DR InterPro; IPR003006; IG_MHC.  
DR InterPro; IPR003596; IG_v.  
DR Pfam; PF07654; Cl-set; 2.  
DR SMART; SM00409; IG; 4.  
DR SMART; SM00407; IGV; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG_LIKE; 4.  
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.  
KW Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 506 AA; 54459 MW; ID9477A37B7C3A0 CRC64;  
  
Query Match 99.5%; Score 1889; DB 2; Length 506;  
Best Local Similarity 99.4%; Pred. No. 1.6e-111;  
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 SPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 61  
Db 168 SPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 227
```

```
Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 60  
Db 154 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 213  
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKYTNPSQDVTPCPVPSPPTPTPTPTPSP 120  
Db 214 GDLYTTSSQLTLPATQCLAGKSVTCHVKYTNPSQDVTPCPVPSPPTPTPTPTPSP 273  
QY 121 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDPDLC 180  
Db 274 SCCHPRLSLHRPALEDLLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDPDLC 333  
QY 181 GCYSVSVSLGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 240  
Db 334 GCYSVSVSLGCAEPNHNHGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSSEEL 393  
QY 241 ALNELVTLTCLARGFSPKDVLRWLOGSQELPREKYLTVASRQEPSQGTTFFAVTSILRV 300  
Db 394 ALNELVTLTCLARGFSPKDVLRWLOGSQELPREKYLTVASRQEPSQGTTFFAVTSILRV 453  
QY 301 AEDWKKGDFTFCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 454 AEDWKKGDFTFCWVGHEALPLAFTOKTIDRLAGKPTHVNVSVVMAEVDGTCY 506  
  
RESULT 12  
Q6N092 PRELIMINARY; PRT; 519 AA.  
ID Q6N092  
AC Q6N092; 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein DKFZp686K18196 (Fragment).  
GN Name=DKFZp686K18196;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Human esophagus tumor;  
RG The German Human cDNA Consortium;  
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,  
RA Fobo G., Han M., Wiemann S.;  
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
EMBL; BX640824; CAB45778.1; -  
DR HSP; P01842; IAGK.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003597; IG-cl.  
DR InterPro; IPR003006; IG_MHC.  
DR InterPro; IPR003596; IG_v.  
DR Pfam; PF07654; Cl-set; 2.  
DR Pfam; PF00047; IG; 1.  
DR SMART; SM00409; IG; 4.  
DR SMART; SM00407; IGV; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG_LIKE; 4.  
DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.  
KW Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 519 AA; 56423 MW; F839BE7F811EB88D CRC64;  
  
Query Match 99.5%; Score 1889; DB 2; Length 519;  
Best Local Similarity 99.7%; Pred. No. 1.6e-111;  
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 SPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 61  
Db 168 SPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNPPPSQDAS 227
```

```

QY 62 DLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTSPSPPTSPS 121
DB 228 DLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTSPSPPTSPS 287
QY 122 CCHPRLSLHRPALEDLLLGSEANLTCITGLRDASGVFTFTWTPSSGKSAVQGPDRDLG 181
DB 288 CCHPRLSLHRPALEDLLLGSEANLTCITGLRDASGVFTFTWTPSSGKSAVQGPDRDLG 347
QY 182 CYSVSVSVLSCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 241
DB 348 CYSVSVLPGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 407
QY 242 LNELVLTLCIARGFSPKDVLRVWLQGSQELPREKYLTVASRQPSQGTTFVAVTSILRV 301
DB 408 LNELVLTLCIARGFSPKDVLRVWLQGSQELPREKYLTVASRQPSQGTTFVAVTSILRV 467
QY 302 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHNVSVVMAEVDGTCY 353
DB 468 AEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHNVSVVMAEVDGTCY 519

RESULT 13
Q6N090
ID C6N090 PRELIMINARY; PRT; 506 AA.
AC C6N090;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKF2p686G21220;
GN Name=DKF2p686G21220;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human rectum tumor;
RG The German Human cDNA Consortium;
RA Wambutt R., Heubner D., News H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640626; CAE45780.1; -.
DR HSSP; P01751; 1A6W.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG.CI.
DR InterPro; IPR003006; IG.MHC.
DR InterPro; IPR003596; IG.V.
DR Pfam; PF07654; C1-set; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IG1; 2.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG.LIKE; 4.
DR PROSITE; PS00290; IG.MHC; UNKNOWN_1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 506 AA; 54388 MW; CCC670F063D32612 CRC64;

Query Match 99.3%; Score 1884; DB 2; Length 506;
Best Local Similarity 99.2%; Pred.No. 3.3e-111;
Matches 350; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGGVTVARNPPSQDAS 60
DB 154 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGGVTVARNPPSQDAS 213
QY 61 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTSP 120
DB 214 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTSP 273
QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCITGLRDASGVFTFTWTPSSGKSAVQGPDRDLG 180
DB 274 SCCHPRLSLHRPALEDLLLGSEANLTCITGLRDASGVFTFTWTPSSGKSAVQGPDRDLG 333

```

```

QY 181 GCYSVSVSVLSCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240
DB 334 GCYSVSVSVLPGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 393
QY 241 ALNELVLTLCIARGFSPKDVLRVWLQGSQELPREKYLTVASRQPSQGTTFVAVTSILRV 300
DB 394 ALNELVLTLCIARGFSPKDVLRVWLQGSQELPREKYLTVASRQPSQGTTFVAVTSILRV 453
QY 301 AADWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHNVSVVMAEVDGTCY 353
DB 454 AADWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHNVSVVMAEVDGTCY 506

RESULT 14
Q8WY24
ID C8WY24 PRELIMINARY; PRT; 497 AA.
AC Q8WY24;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SNC66 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng S., Shao X., Cao J., Geng L., Fang Y., Dong Q.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF283666; AAL36987.1; -.
DR HSSP; P01876; 1OW0.
DR Pfam; PF07654; C1-set; 2.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG.LIKE; 4.
DR PROSITE; PS00290; IG.MHC; UNKNOWN_1.
SQ SEQUENCE 497 AA; 53665 MW; F21D08DFA5A663E5 CRC64;

Query Match 99.1%; Score 1881; DB 2; Length 497;
Best Local Similarity 98.9%; Pred.No. 5e-111;
Matches 349; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGGVTVARNPPSQDAS 60
DB 145 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQBPPLSVTWSESGGVTVARNPPSQDAS 204
QY 61 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTSP 120
DB 205 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPCEVPSTPTPTSP 264
QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCITGLRDASGVFTFTWTPSSGKSAVQGPDRDLG 180
DB 265 SCCHPRLSLHRPALEDLLLGSEANLTCITGLRDASGVFTFTWTPSSGKSAVQGPDRDLG 324
QY 181 GCYSVSVSVLSCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240
DB 325 GCYSVSVSVLPGCAEPNHNKGTFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 384
QY 241 ALNELVLTLCIARGFSPKDVLRVWLQGSQELPREKYLTVASRQPSQGTTFVAVTSILRV 300
DB 385 ALNELVLTLCIARGFSPKDVLRVWLQGSQELPREKYLTVASRQPSQGTTFVAVTSILRV 444
QY 301 AADWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHNVSVVMAEVDGTCY 353
DB 445 AADWKKGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHNVSVVMAEVDGTCY 497

RESULT 15
ALCL_GORGO
ID ALCL_GORGO STANDARD; PRT; 353 AA.
AC P20758;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)

```

[illegible]

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:02:42 ; Search time 25.1544 Seconds  
(without alignments)  
1047.575 Million cell updates/sec

Title: US-10-644-256-3\_COPY\_137\_489

Perfect score: 1898

Sequence: 1 ASPTSPKVFPLSLCSTQPDG.....GKETHVNVSVMAEVDGTCY 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1867	98.4	530	4	US-09-800-729-112
2	1862	98.1	494	4	US-09-800-729-216
3	1861	98.1	499	3	US-09-049-672A-1
4	1710.5	90.1	483	3	US-09-049-672A-5
5	1710.5	90.1	487	4	US-09-800-729-145
6	1267.5	66.8	343	1	US-08-336-891-2
7	1267.5	66.8	343	5	PCT-US95-13795-4
8	585	30.8	113	1	US-08-399-106A-4
9	585	30.8	113	1	US-08-433-105A-4
10	585	30.8	113	2	US-08-434-869A-4
11	505	26.6	504	1	US-07-932-915-2
12	505	26.6	504	5	PCT-US91-05826-2
13	444.5	23.4	362	4	US-09-291-299A-6
14	444.5	23.4	470	4	US-09-291-299A-1
15	444.5	23.4	476	4	US-09-291-299A-3
16	442	23.3	375	4	US-09-828-995B-32
17	441	23.2	334	2	US-08-646-981-16
18	439	23.1	333	1	US-08-436-463-6
19	439	23.1	333	1	US-08-024-253-6
20	438	23.1	473	4	US-09-828-995B-20
21	421	22.2	470	4	US-09-828-995B-11
22	416	21.9	377	2	US-08-761-277A-4
23	415	21.9	352	4	US-09-828-995B-38
24	409.5	21.6	552	5	PCT-US93-07832-23
25	406.5	21.4	331	2	US-08-646-981-17
26	404.5	21.3	451	4	US-09-472-087-70
27	404.5	21.3	463	4	US-09-472-087-1

28	404.5	21.3	463	4	US-09-472-087-4
29	404.5	21.3	463	4	US-09-472-087-63
30	404.5	21.3	463	4	US-09-472-087-64
31	404.5	21.3	463	4	US-09-472-087-68
32	404.5	21.3	464	4	US-09-472-087-2
33	404.5	21.3	464	4	US-09-472-087-66
34	404.5	21.3	470	4	US-09-859-053-28
35	404.5	21.3	470	4	US-09-859-053-32
36	404.5	21.3	470	4	US-09-859-053-36
37	404.5	21.3	530	3	US-08-477-460B-4
38	404.5	21.3	530	3	US-08-379-916-4
39	404.5	21.3	530	3	US-09-329-916-4
40	404.5	21.3	530	3	US-08-485-372A-4
41	404.5	21.3	530	3	US-09-409-006A-4
42	404.5	21.3	530	4	US-08-484-681-4
43	404.5	21.3	530	4	US-09-766-995-4
44	404.5	21.3	530	5	PCT-US93-07422-4
45	404	21.3	468	4	US-09-828-995B-5

## ALIGNMENTS

RESULT 1  
US-09-800-729-112  
; Sequence 112, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: PZ044P1  
; CURRENT APPLICATION NUMBER: US/09/800, 729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 112  
; LENGTH: 530  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (488)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (490)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (494)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (495)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (505)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-800-729-112

Query Match 98.4%; Score 1867; DB 4; Length 530;

Best Local Similarity 98.9%; Pred. No. 2e-143;

Matches 348; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	1	ASPTSPKVFPLSLCSTQPDGNNVTACLVQGFQFQEPUSVTWSESGQGVTAARNFPSPQDAS 60
Db	142	ASPTSPKVFPLSLCSTQPDGNNVTACLVQGFQFQEPUSVTWSESGQGVTAARNFPSPQDAS 201
Qy	61	GDLYTTSSQLTLPATQCLACKSVTCHKVHTNPSQDVTVPSPVSTPTPTSP 120
Db	202	GDLYTTSSQLTLPATQCLACKSVTCHKVHTNPSQDVTVPSPVSTPTPTSP 261

QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTWTPSSGKSAVQPPDRDLC 180  
Db 262 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTWTPSSGKSAVQPPDRDLC 321  
QY 181 GCYSVSVSLGCAEPNNHGKFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Db 322 GCYSVSVSLPGCAEPNNHGKFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 381  
QY 241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFVAVTSILRV 300  
Db 382 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFVAVTSILRV 441  
QY 301 AADWKKGTFTFCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTC 352  
Db 442 AADWKKGTFTFCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTC 493

## RESULT 2

US-09-800-729-216  
; Sequence 216, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800, 729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCI/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155, 709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 216  
; LENGTH: 494  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-216

Query Match 98.1%; Score 1862; DB 4; Length 494;  
Best Local Similarity 97.5%; Pred. No. 4.6e-143;  
Matches 344; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60  
Db 142 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 201  
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPTPTPTPTPTPTPTSP 120  
Db 202 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPTPTPTPTPTPTPTSP 261  
QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTWTPSSGKSAVQPPDRDLC 180  
Db 262 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTWTPSSGKSAVQPPDRDLC 321  
QY 181 GCYSVSVSLGCAEPNNHGKFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Db 322 GCYSVSVSLPGCAEPNNHGKFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 381  
QY 241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFVAVTSILRV 300  
Db 382 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFVAVTSILRV 441  
QY 301 AADWKKGTFTFCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTC 353  
Db 442 AADWKKGTFTFCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTC 494

## RESULT 3

US-09-049-672A-1  
; Sequence 1, Application US/09049672A  
; Patent No. 6135941  
; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Yue, Henry  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/049,672A  
; FILING DATE: HERewith  
; CLASSIFICATION: 536  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cerrone, Michael C  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0497 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 499 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: ADENINB01  
; CLONE: 021145  
; US-09-049-672A-1

Query Match 98.1%; Score 1861; DB 3; Length 499;  
Best Local Similarity 97.5%; Pred. No. 5.7e-143;  
Matches 344; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60  
Db 147 ASPTSPKVPFLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 206  
QY 61 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPTPTPTPTPTPTPTSP 120  
Db 207 GDLYTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVPTPTPTPTPTPTPTSP 266  
QY 121 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTWTPSSGKSAVQPPDRDLC 180  
Db 267 SCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVFTWTPSSGKSAVQPPDRDLC 326  
QY 181 GCYSVSVSLGCAEPNNHGKFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Db 327 GCYSVSVSLPGCAEPNNHGKFTCTAAHPPELTPLTANITKSGNTRFRPEVHLLPPPEEL 386  
QY 241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFVAVTSILRV 300  
Db 387 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVASRQEPSQGTTFVAVTSILRV 446  
QY 301 AADWKKGTFTFCMWGHEALPLAFTQKTDRLAGKPTHVNVSVVMAEVDGTC 353

Db 447 AADWKKGGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 499  
|||||

## RESULT 4

US-09-049-672A-5  
; Sequence 5, Application US/09049672A  
; Patent No. 6135941  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Yue, Henry  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/049,672A  
; FILING DATE: HERewith  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cerrone, Michael C  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PP-0497 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 483 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BMARNO703  
; CLONE: 1669829

US-09-049-672A-5

Query Match 90.1%; Score 1710.5; DB 3; Length 483;  
Best Local Similarity 90.9%; Pred. No. 9e-131;  
Matches 321; Conservative 7; Mismatches 12; Indels 13; Gaps 1;  
  
Qy 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60  
Db 144 ASPTSPKVFPLSLDSTPDQGNVVVACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 203  
  
Qy 61 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPSP 120  
Db 204 GDLYTSSQLTLPATQCPDGKSVTCHVKHYTNPSQDVTVPVSTPTPTPSP 250  
  
Qy 121 SCCHPRLSLHRPALEDLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 180  
Db 251 PCCHPRLSLHRPALEDLLGSEANLCTLTGLRDASGATFTWTPSSGKSAVQGPDRDLC 310

Qy 181 GCYSVSVSLSGCAEPWNHKGKTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Db 311 GCYSVSVSLPGCAQPNWNGHGETTCTAAHPKLTPLTANITKSGNTFRPEVHLLPPPEEL 370  
  
Qy 241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLWTASRQPSQGTTFVAVTSILRV 300  
Db 371 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLWTASRQPSQGTTFVAVTSILRV 430  
  
Qy 301 AADWKKGGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 431 AADWKKGGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 483

## RESULT 5

US-09-800-729-145  
; Sequence 145, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: PZ044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 145  
; LENGTH: 487  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-145

Query Match 90.1%; Score 1710.5; DB 4; Length 487;  
Best Local Similarity 90.9%; Pred. No. 9.1e-131;  
Matches 321; Conservative 7; Mismatches 12; Indels 13; Gaps 1;

Qy 1 ASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 60  
Db 148 ASPTSPKVFPLSLDSTPDQGNVVVACLVQGFPPQEPPLSVTWSESGQGVTAARNFPPSQDAS 207  
  
Qy 61 GDLYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPVSTPTPTPSP 120  
Db 208 GDLYTSSQLTLPATQCPDGKSVTCHVKHYTNPSQDVTVPVSTPTPTPSP 254  
  
Qy 121 SCCHPRLSLHRPALEDLLGSEANLCTLTGLRDASGVTFTWTPSSGKSAVQGPDRDLC 180  
Db 255 PCCHPRLSLHRPALEDLLGSEANLCTLTGLRDASGATFTWTPSSGKSAVQGPDRDLC 314  
  
Qy 181 GCYSVSVSLSGCAEPWNHKGKTTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPEEL 240  
Db 315 GCYSVSVSLPGCAQPNWNGHGETTCTAAHPKLTPLTANITKSGNTFRPEVHLLPPPEEL 374  
  
Qy 241 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLWTASRQPSQGTTFVAVTSILRV 300  
Db 375 ALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLWTASRQPSQGTTFVAVTSILRV 434  
  
Qy 301 AADWKKGGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 353  
Db 435 AADWKKGGDTFSCWVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY 487

## RESULT 6

US-08-336-891-2  
; Sequence 2, Application US/08336891  
; Patent No. 5622842  
; GENERAL INFORMATION:  
; APPLICANT: HOLLIS, GREGORY F.  
; APPLICANT: PATEL, MAYUR D.  
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN A  
; NUMBER OF SEQUENCES: 2



```
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,106A
; FILING DATE: 01-Mar-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: P0927D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-399-106A-4

Query Match 30.8%; Score 585; DB 1; Length 113;
Best Local Similarity 96.5%; Pred. No. 3.2e-40;
Matches 109; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 223 GNTFRPEVHLLPPPSBELALNELVTLTCLARGSPKDVLRVWLGSOELPREKYLTVASR 282
Db 1 GNTFRPQVHLLPPPSBELALBZLVTLTCLARGSPKDVLRVWLGSOELPREKYLTVASR 60

Qy 283 QEPSQGTTFVAVTSILRVAEDWKGGDTFCMVGHEALPLAFTQKTDRLAGK 335
Db 61 QZPSQGTTFVAVTSILRVAEDWKGGDTFCMVGHEALPLAFTQKTDRLAGK 113

RESULT 9
US-08-433-105A-4
; Sequence 4, Application US/08433105A
; Patent No. 5807706
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Ridgway, John B.
; TITLE OF INVENTION: A METHOD FOR MAKING HETEROMULTIMERIC POLYPEPTIDES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,105A
; FILING DATE: 03-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/399106
; FILING DATE: 01-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: P0927D2
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-433-105A-4

Query Match 30.8%; Score 585; DB 1; Length 113;
Best Local Similarity 96.5%; Pred. No. 3.2e-40;
Matches 109; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 223 GNTFRPEVHLLPPPSBELALNELVTLTCLARGSPKDVLRVWLGSOELPREKYLTVASR 282
Db 1 GNTFRPQVHLLPPPSBELALBZLVTLTCLARGSPKDVLRVWLGSOELPREKYLTVASR 60

Qy 283 QEPSQGTTFVAVTSILRVAEDWKGGDTFCMVGHEALPLAFTQKTDRLAGK 335
Db 61 QZPSQGTTFVAVTSILRVAEDWKGGDTFCMVGHEALPLAFTQKTDRLAGK 113

RESULT 10
US-08-434-869A-4
; Sequence 4, Application US/08434869A
; Patent No. 5821333
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Ridgway, John B.
; TITLE OF INVENTION: A METHOD FOR MAKING HETEROMULTIMERIC POLYPEPTIDES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,869A
; FILING DATE: 03-May-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/399106
; FILING DATE: 01-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: P0927D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-434-869A-4

Query Match 30.8%; Score 585; DB 2; Length 113;
Best Local Similarity 96.5%; Pred. No. 3.2e-40;
Matches 109; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 223 GNTFRPEVHLLPPPSBELALNELVTLTCLARGSPKDVLRVWLGSOELPREKYLTVASR 282
```

Db 1 GNTFRPQVHLLPPPSBELABZLVLTCLTCLARGFSPKDVLRVWLQSGQELPREKYLTVASR 60

QY 283 QEPSQGTTFVAVTSILRVAEDWKKGDTSCMGVHGHEALPLAFTQKIDRLAGK 335

Db 61 QZPSQGTTFVAVTSILRVAEDWKKGDTSCMGVHGHEALPLAFTQKIDRLAGK 113

RESULT 11

US-07-932-915-2

; Sequence 2, Application US/07932915

; Patent No. 5672486

; GENERAL INFORMATION:

; APPLICANT: Soullilou, Jean-Paul

; TITLE OF INVENTION: Protein Polyligands Joined To A Stable Protein

; TITLE OF INVENTION: Core

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum

; STREET: 5 Palo Alto Square, Suite 400

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/932,915

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/646,875

; FILING DATE: 28-JAN-1991

; APPLICATION NUMBER: US 07/575,394

; FILING DATE: 23-AUG-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland Ph.D., Bertram I.

; REGISTRATION NUMBER: 20,015

; REFERENCE/DOCKET NUMBER: ATLA-001/01US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-494-7622

; TELEFAX: 415-857-0663

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 504 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-932-915-2

Query Match 26.6%; Score 505; DB 1; Length 504;

Best Local Similarity 33.3%; Pred. No. 7.2e-33;

Matches 118; Conservative 49; Mismatches 131; Indels 56; Gaps 9;

QY 24 IACLVOGFFPQELSVTW-----SESGQGVTAARNFPSPQDASG-DLYTSSQLTLPATQCL 78

Db 183 LICQATGFSRQ-IQVSWLRGKQVGSVTTDOVQAEAKESGPTTKYKVTSTLIKESDWL 241

QY 79 AGKSVTCHVKH-----YTNPSQDVTPCPVPSTPTPTSPSTPTPTSPSCCHPRL 127

Db 242 SQSMFTCRVDHRLGLTQQNASSMCPDQDTAI-----RV 275

QY 128 SLHRPALEDLLGSEANLCTLTGLRDASGVFTTTPSSGKSA-----VQGPDRDLGC 181

Db 276 FAIPPSFASIFLTKSKLTCLVTLDTYDSVTSIWMTRQNGEAVKTHNISSEHPN----A 331

QY 182 CYSVSVSLGCAEPNHNKGTFTCTAAYPESKTPLTATLS--KSGNTFRPEVHLPPPSEE 239

Db 332 TFSAVGEASICEDDMNSGERFTCTVTHDLPSPKQTSIRPKGVALLHRDVLVLPAREQ 391

QY 240 LALNELVTLTCLARGFSPKDVLRVWLQSGQELPREKYLTVASRQEPSQGTTFVAVTSILR 299

Db 392 LNLRESATITCLVTGFSPADVFQVMQRCQPLSPKDYVTSAPMPEP-QAPGRYFAHSILT 450

QY 300 VAAEDWKKGDTSCMGVHGHEALPLAFTQKIDRLAGKPTHVNVSVVMAEVDGTCY 353

Db 451 VSEEWNTGETVTCVVAHEALPNRVTVKSTGKPTLYNVSLVMSDAGTCY 504

RESULT 12

PCT-US91-05826-2

; Sequence 2, Application PC/TUS9105826

; GENERAL INFORMATION:

; APPLICANT: Soullilou, Jean-Paul

; TITLE OF INVENTION: Protein Polyligands Joined To A Stable Protein

; TITLE OF INVENTION: Core

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bertram I. Rowland, Ph.D.

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US91/05826

; FILING DATE: 19910822

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/575,394

; FILING DATE: 29-AUG-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland Ph.D., Bertram I.

; REGISTRATION NUMBER: 20,015

; REFERENCE/DOCKET NUMBER: FP55352-1/BIR

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-494-8771

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 504 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US91-05826-2

Query Match 26.6%; Score 505; DB 5; Length 504;

Best Local Similarity 33.3%; Pred. No. 7.2e-33;

Matches 118; Conservative 49; Mismatches 131; Indels 56; Gaps 9;

QY 24 IACLVOGFFPQELSVTW-----SESGQGVTAARNFPSPQDASG-DLYTSSQLTLPATQCL 78

Db 183 LICQATGFSRQ-IQVSWLRGKQVGSVTTDOVQAEAKESGPTTKYKVTSTLIKESDWL 241

QY 79 AGKSVTCHVKH-----YTNPSQDVTPCPVPSTPTPTSPSTPTPTSPSCCHPRL 127

Db 242 SQSMFTCRVDHRLGLTQQNASSMCPDQDTAI-----RV 275

QY 128 SLHRPALEDLLGSEANLCTLTGLRDASGVFTTTPSSGKSA-----VQGPDRDLGC 181

Db 276 FAIPPSFASIFLTKSKLTCLVTLDTYDSVTSIWMTRQNGEAVKTHNISSEHPN----A 331

QY 182 CYSVSVSLGCAEPNHNKGTFTCTAAYPESKTPLTATLS--KSGNTFRPEVHLPPPSEE 239

Db 332 TFSAVGEASICEDDMNSGERFTCTVTHDLPSPKQTSIRPKGVALLHRDVLVLPAREQ 391

QY 240 LALNELVTLTCLARGFSPKDVLRVWLQSGQELPREKYLTVASRQEPSQGTTFVAVTSILR 299

Db 392 LNLRESATITCLVTGFSPADVFQVMQRCQPLSPKDYVTSAPMPEP-QAPGRYFAHSILT 450



Search completed: October 25, 2005, 11:09:32  
Job time : 26.1544 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:07:58 ; Search time 131.252 Seconds  
(without alignments)  
1555.504 Million cell updates/sec

Title: US-10-644-256-3

Perfect score: 2611

Sequence: 1 MACPGFLWALVISTCLEFSM.....GKPTHVNVSVVMAEVDGCTY 489

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1862994 seqs, 417510619 residues

Total number of hits satisfying chosen parameters: 1862994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2611	100.0	489	17	US-10-644-256-3
2	2335	89.4	630	15	Sequence 3, Appl
3	2335	89.4	630	15	Sequence 48, Appl
4	2302	88.2	496	15	Sequence 16, Appl
5	2281.5	87.4	496	15	Sequence 4058, Ap
6	2276	87.2	495	15	Sequence 3006, Ap
7	2275	87.1	502	15	Sequence 4085, Ap
8	2275	87.1	495	15	Sequence 4245, Ap
9	2215	84.8	494	15	Sequence 4277, Ap
10	2200.5	84.3	497	15	Sequence 798, App
11	2192	84.0	494	15	Sequence 3773, Ap
					Sequence 4078, Ap

12	2190	83.9	508	15	US-10-104-047-3233	Sequence 3233, Ap
13	2188.5	83.8	497	15	US-10-108-260A-4244	Sequence 4244, Ap
14	2186	83.7	530	9	US-09-800-729-112	Sequence 112, App
15	2186	83.7	530	11	US-09-833-245-2189	Sequence 2189, Ap
16	2181	83.5	494	9	US-09-800-729-216	Sequence 216, App
17	2179.5	83.5	491	15	US-10-108-260A-4262	Sequence 4262, Ap
18	2174	83.3	500	15	US-10-108-260A-4255	Sequence 4255, Ap
19	2172.5	83.2	491	15	US-10-104-047-3243	Sequence 3243, Ap
20	2169	83.1	494	15	US-10-108-260A-4275	Sequence 4275, Ap
21	2160.5	82.7	495	15	US-10-108-260A-4114	Sequence 4114, Ap
22	2125.5	81.4	491	15	US-10-108-260A-4290	Sequence 4290, Ap
23	2120.5	81.2	495	11	US-09-833-245-302	Sequence 302, App
24	2120	81.2	686	15	US-10-072-012-795	Sequence 795, App
25	2113.5	80.9	475	15	US-10-072-012-794	Sequence 794, App
26	2110	80.8	472	15	US-10-072-012-796	Sequence 796, App
27	2100	80.4	508	15	US-10-108-260A-3028	Sequence 3028, Ap
28	2092.5	80.1	500	15	US-10-108-260A-4684	Sequence 4684, Ap
29	2025.5	77.6	669	9	US-09-807-721-2	Sequence 2, Appl
30	1994.5	76.4	393	14	US-10-221-945-3	Sequence 3, Appl
31	1970.5	75.5	487	9	US-09-800-729-145	Sequence 145, App
32	1970.5	75.5	487	11	US-09-833-245-2194	Sequence 2194, Ap
33	1969	75.4	384	15	US-10-072-012-797	Sequence 797, App
34	1889	72.3	353	13	US-10-047-542-16	Sequence 16, Appl
35	1888	72.3	353	20	US-11-003-819-55	Sequence 55, Appl
36	1878.5	71.9	354	17	US-10-872-932A-32	Sequence 32, Appl
37	1878.5	71.9	354	18	US-10-810-881A-31	Sequence 31, Appl
38	1710.5	65.5	340	13	US-10-047-542-18	Sequence 18, Appl
39	1710.5	65.5	340	17	US-10-872-932A-33	Sequence 33, Appl
40	1710.5	65.5	340	18	US-10-810-881A-32	Sequence 32, Appl
41	1695.5	64.9	538	13	US-10-047-542-99	Sequence 99, Appl
42	1684	64.5	799	13	US-10-047-542-8	Sequence 8, Appl
43	1684	64.5	822	13	US-10-047-542-48	Sequence 48, Appl
44	1679	64.3	338	15	US-10-072-012-322	Sequence 322, App
45	1640	62.8	521	18	US-10-627-556-398	Sequence 398, App

#### ALIGNMENTS

RESULT 1  
US-10-644-256-3  
; Sequence 3, Application US/10644256  
; Publication No. US20050106722A1  
; GENERAL INFORMATION:  
; APPLICANT: Jones, David HA  
; APPLICANT: Bout, Abraham  
; TITLE OF INVENTION: Efficient Production of IgA in Recombinant Mammalian Cells  
; FILE REFERENCE: 2578-6077  
; CURRENT APPLICATION NUMBER: US/10/644,256  
; CURRENT FILING DATE: 2003-08-20  
; PRIOR APPLICATION NUMBER: US 09/549,463  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: US 60/129,452  
; PRIOR FILING DATE: 1999-04-15  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: artificial  
; FEATURE:  
; OTHER INFORMATION: Amino acid sequence anti-EpCAM IgA heavy chain  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(21)  
; OTHER INFORMATION: leader peptide  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (22)..(136)  
; OTHER INFORMATION: VH region  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (137)..(238)

; OTHER INFORMATION: CH1 Region  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (239]..(359)  
; OTHER INFORMATION: CH2 Region  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (360]..(489)  
; OTHER INFORMATION: CH3 Region  
US-10-644-256-3

Query Match 100.0%; Score 2611; DB 17; Length 489;  
Best Local Similarity 100.0%; Pred.No. 7.5e-147;  
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPGFLWALVISTCLFSEFSAQVQLVQSGAEVKKPGSSVRVSKASGGTFSSVAISWVRQ 60  
DB 1 MACPGFLWALVISTCLFSEFSAQVQLVQSGAEVKKPGSSVRVSKASGGTFSSVAISWVRQ 60  
QY 61 APGGLEWMMGIIPIFGTANYAOKFQGRVTITADESTAYMELSLRSRSEDYAVYCARD 120  
DB 61 APGGLEWMMGIIPIFGTANYAOKFQGRVTITADESTAYMELSLRSRSEDYAVYCARD 120  
QY 121 PFLHYWQGLTVTVSTASPTSPKVFPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWS 180  
DB 121 PFLHYWQGLTVTVSTASPTSPKVFPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWS 180  
QY 181 GQGVARNFPPSDASDGLYTSQSLLTPATQCLAGKSVTCHVKHTNPSQDVTVPVCP 240  
DB 181 GQGVARNFPPSDASDGLYTSQSLLTPATQCLAGKSVTCHVKHTNPSQDVTVPVCP 240  
QY 241 STPTSPSPPTPPSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTTTWP 300  
DB 241 STPTSPSPPTPPSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTTTWP 300  
QY 301 SSGKSAVQGPDPDRDLCCYSSVSVLSCAEPNHNKFTTCTAAYPESKPTLTATLSKSGN 360  
DB 301 SSGKSAVQGPDPDRDLCCYSSVSVLSCAEPNHNKFTTCTAAYPESKPTLTATLSKSGN 360  
QY 361 TFRPEVHLLPPSEELALNELVLTCLARGFSPKDVLRWLGQSOELPREKYLTVASRQE 420  
DB 361 TFRPEVHLLPPSEELALNELVLTCLARGFSPKDVLRWLGQSOELPREKYLTVASRQE 420  
QY 421 PSQGTITFAVTSILRVAEDWKKGDTFSCVMGHEALPLAFTQKTIIDRLAGKPTHVNVSVV 480  
DB 421 PSQGTITFAVTSILRVAEDWKKGDTFSCVMGHEALPLAFTQKTIIDRLAGKPTHVNVSVV 480  
QY 481 MAEVDGTCY 489  
DB 481 MAEVDGTCY 489

RESULT 2  
US-10-422-628-48  
; Sequence 48, Application US/10422628  
; Publication No. US20040014174A1  
; GENERAL INFORMATION:  
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE  
; APPLICANT: MAYFIELD, Stephen P.  
; TITLE OF INVENTION: EXPRESSION OF POLYPEPTIDES IN CHLOROPLASTS, AND COMPOSITIONS AND  
; FILE REFERENCE: SCRIPI510-2  
; CURRENT APPLICATION NUMBER: US 60/434,957  
; PRIOR FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/375,129  
; PRIOR FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/434,957  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 48  
; LENGTH: 630

; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Single-chain antibody  
US-10-422-628-48

Query Match 89.4%; Score 2335; DB 15; Length 630;  
Best Local Similarity 92.4%; Pred.No. 2.3e-130;  
Matches 439; Conservative 11; Mismatches 15; Indels 10; Gaps 1;

QY 25 LVQSGAEVKKPGSSVRVSKASGGTFSSVAISWVRQAPQGLEWMMGIIPIFGTANYAOK 84  
DB 134 LEQSGAEVKKPGSSVRVSKASGGTFSSVAISWVRQAPQGLEWMMGIIPIFGTANYAOK 193  
QY 85 FQGRVTITADESTAYMELSLRSRSEDYAVYCARDPF-----LHYWQGLTVTV 134  
DB 194 FQGRVTITADESTAYMELSLRSRSEDYAVYCARDPF-----LHYWQGLTVTV 253  
QY 135 STASPTSPKVFPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSGQGVARNFPPSQD 194  
DB 254 STASPTSPKVFPPLSLCSTQPDGNVVIACLVQGFPPQEPPLSVTWSGQGVARNFPPSQD 313  
QY 195 ASGDLTYTSQSLLTPATQCLAGKSVTCHVKHTNPSQDVTVPVCPSTPTPTPTPTPT 254  
DB 314 ASGDLTYTSQSLLTPATQCLAGKSVTCHVKHTNPSQDVTVPVCPSTPTPTPTPTPT 373  
QY 255 SPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTTTWPSSGKSAVQGPDPDR 314  
DB 374 SPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGLRDASGVTTTWPSSGKSAVQGPDPDR 433  
QY 315 LCCYSSVSVLSCAEPNHNKFTTCTAAYPESKPTLTATLSKSGNTFRPEVHLLPPSPSE 374  
DB 434 LCCYSSVSVLSCAEPNHNKFTTCTAAYPESKPTLTATLSKSGNTFRPEVHLLPPSPSE 493  
QY 375 ELALNELVLTCLARGFSPKDVLRWLGQSOELPREKYLTVASRQPSQGTITFAVTSIL 434  
DB 494 ELALNELVLTCLARGFSPKDVLRWLGQSOELPREKYLTVASRQPSQGTITFAVTSIL 553  
QY 435 RVAEDWKKGDTFSCVMGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 489  
DB 554 RVAEDWKKGDTFSCVMGHEALPLAFTQKTIIDRLAGKPTHVNVSVVMAEVDGTCY 608

RESULT 3  
US-10-422-628-16  
; Sequence 16, Application US/10422628  
; Publication No. US20040014174A1  
; GENERAL INFORMATION:  
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE  
; APPLICANT: MAYFIELD, Stephen P.  
; APPLICANT: FRANKLIN, Scott E.  
; TITLE OF INVENTION: EXPRESSION OF POLYPEPTIDES IN CHLOROPLASTS, AND COMPOSITIONS AND  
; FILE REFERENCE: SCRIPI510-2  
; CURRENT APPLICATION NUMBER: US 60/434,957  
; PRIOR FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/375,129  
; PRIOR FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/434,957  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 16  
; LENGTH: 639  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Chloroplast codon optimized antibody specific for Herpes simplex  
; OTHER INFORMATION: virus  
US-10-422-628-16

Query Match 89.4%; Score 2335; DB 15; Length 639;  
Best Local Similarity 92.4%; Pred.No. 2.3e-130;

Query Match	87.4%;	Score	2281.5;	DB	15;	Length	496;
Best Local Similarity	90.0%;	Pred. No.	2.6e-127;				
Matches	430;	Conservative	13;	Mismatches	26;	Indels	9; Gaps 1;

  

Qy	21	AQVLQVSGAEVKPKGSSVVRVSCAKSGGTFSSVAISWVRQAPGOGLEWMGGLIPIFGTAN	80
Db	19	SBQQLVQSGAEVKRPGASVKITKACQHTFSSHYHWVRQAPGQGLEWLGIIINPRGGQTT	78
Qy	81	YAKQFGQVRVITADESTSTAYMELSSLRSBDDTAVYYCARDPF-----LHYWGQGTLL	131
Db	79	YAKQLQGRVTMTSDTSTATVMELSSLRSBDDTAVYYCVSAAYCGDCYPGIFDLWGRGTL	138
Qy	132	VTVSTASPTSPKVPFLSLCSTQPDGNVVIACLVOGFEPQEPPLSVTWSESQGVYARNFPP	191
Db	139	VTVSAASPTSPKVPFLSLCSTQPDGNVVIACLVOGFEPQEPPLSVTWSESQGVYARNFPP	198
Qy	192	SDQASGDLYTTSSQLTLPTATQCLAGKAGSVTCHVVKHYTNPSQDVVTPCPVPSTPTPSSTP	251
Db	199	SDQASGDLYTTSSQLTLPTATQCLAGKAGSVTCHVVKHYTNPSQDVVTPCPVPSTPTPSSTP	258
Qy	252	PTSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFWTWTPSSGKSAVOGPP	311
Db	259	PTSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGLRDASGVTFWTWTPSSGKSAVOGPP	318
Qy	312	DRDLCCGYSVSSVLSGCAEPWNHGKFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPP	371
Db	319	ERDLCCGYSVSSVLPCCAEPNWNGKFTCTAAYPESKTPLTATLSKSGNTFRPEVHLLPP	378
Qy	372	PSBELALNELVTLTCLARGFSPKDXLVRLMQGSOELPREKYLVTWASRQEPSQGTTFEAVT	431
Db	379	PSBELALNELVTLTCLARGFSPKDXLVRLMQGSOELPREKYLVTWASRQEPSQGTTFEAVT	438



Db 79 FPPEFLDRVILRRDTSASTVYMEASLTSSEDSALYYCARGTPLRMETEFAYDVMGQSTRV 138  
Qy 133 TVSTASPTSPKVFPLSLCSTOPDGNVVIACLVQGFQFQEPPLSVTWSSGOGVTARNPPS 192  
Db 139 TVSTSTSPKVFPLSLCSTOPDGNVVIACLVQGFQFQEPPLSVTWSSGOGVTARNPPS 198  
Qy 193 QDASGDLYTSSQTLTCLATQCLAGSVTCHVKHYTNPSQDVTVPVPCVPSTPTPTSPSTPP 252  
Db 199 QDASGDLYTSSQTLTCLATQCLAGSVTCHVKHYTNPSQDVTVPVPCVPSTPTPTSPSTPP 258  
Qy 253 TPSPSCCHPRLSLRPALEDLLGSEANLCTLTGLRDASGVTFWTTPSSGKSAVQGPDP 312  
Db 259 TPSPSCCHPRLSLRPALEDLLGSEANLCTLTGLRDASGVTFWTTPSSGKSAVQGPPE 318  
Qy 313 RDLGCGYSVSVSLSCGCAEPNHNHGTFTCTAAYPSKTPLTATLSKSGNTFRPEVHLLPPP 372  
Db 319 RDLGCGYSVSVSLSCGCAEPNHNHGTFTCTAAYPSKTPLTATLSKSGNTFRPEVHLLPPP 378  
Qy 373 SEELALNELVTLTCLARGFSPKDVLRWLQGSQBLPREKYLTVASROEPSQGTTFPAVTS 432  
Db 379 SEELALNELVTLTCLARGFSPKDVLRWLQGSQBLPREKYLTVASROEPSQGTTFPAVTS 438  
Qy 433 ILRVAEDWKKGDTFSCVMGHEALPLAFTQKTIDRLACKPHTHNVSVVMAEVDGTCY 489  
Db 439 ILRVAEDWKKGDTFSCVMGHEALPLAFTQKTIDRLACKPHTHNVSVVMAEVDGTCY 495

RESULT 9

US-10-072-012-798  
; Sequence 798, Application US/10072012  
; Publication No. US20040033493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zethusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsobrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,412  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,395  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/266,406  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/266,767

; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: 60/267,057  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/266,975  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/267,459  
; PRIOR FILING DATE: 2001-02-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1391  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 798  
; LENGTH: 494  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-072-012-798  
  
Query Match 84.8%; Score 2215; DB 15; Length 494;  
Best Local Similarity 84.8%; Pred.No. 2.3e-123; Indels 8; Gaps 2;  
Matches 417; Conservative 26; Mismatches 41;  
  
Qy 5 GFLMALVISTCLBFSMAQVOLVQSGAEVKPGSSVRVSKASGGTTFSSYAIISWVRQAPGQ 64  
Db 4 GLRWVFLVA-FLGVQCEVOLVESGGGLVKPGGSLRLSCAASGLSFSTYAMNWRQAPGK 62  
Qy 65 GLEWMGGIIPFICTANYAQKFGQRTVTADSTSTAYMELSSLRSEDTAVYYCARDPFF-- 122  
Db 63 GLEWVSSISRSYIYYRDSVKGRFTISRDNKNSLYLQMNLSLRVDDTAVYYCARDSCNG 122  
Qy 123 ----LHYWGQGTLVTVSTASPTSPKVPFLSLCSTOPDGNVVIACLVQGFQFQEPPLSVTW 177  
Db 123 AICYGFSPWQGGTLVTVSSASPTSPKVPFLSLCSTQPDGNVVIACLVQGFQFQEPPLSVTW 182  
Qy 178 SESGQGVTAARNFPPSPQDASGDLTYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVP 237  
Db 183 SESGQGVTAARNFPPSPQDASGDLTYTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVP 242  
Qy 238 PVPSTPT 297  
Db 243 PVPSTPT 302  
Qy 298 WTPSSGKSAVQGPDRDLCCGYSVSVSLSCGCAEPNHNHGTFTCTAAYPSKTPLTATLSK 357  
Db 303 WTPSSGKSAVQGPDRDLCCGYSVSVSLSCGCAEPNHNHGTFTCTAAYPSKTPLTATLSK 362  
Qy 358 SGNTFRPEVHLLPPSPSEALNELVTLTCLARGFSPKDVLRWLQGSQBLPREKYLTVAS 417  
Db 363 SGNTFRPEVHLLPPSPSEALNELVTLTCLARGFSPKDVLRWLQGSQBLPREKYLTVAS 422  
Qy 418 ROEPSQGTTFPAVTSILRVAEDWKKGDTFSCVMGHEALPLAFTQKTIDRLACKPHTHNV 477  
Db 423 ROEPSQGTTFPAVTSILRVAEDWKKGDTFSCVMGHEALPLAFTQKTIDRLACKPHTHNV 482  
Qy 478 SVVMAEVDGTCY 489  
Db 483 SVVMAEVDGTCY 494

RESULT 10

US-10-104-047-3773  
; Sequence 3773, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA  
; FILE REFERENCE: HI-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3773  
; LENGTH: 497





```
Db 363 SGNTRPEVHLLPPSEELALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTVAS 422
QY 418 ROEPSGTTTFAVTSILRVAEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNV 477
Db 423 ROEPSGTTTFAVTSILRVAEDWKKGDTFSCWVGHEALPLAFTOKTIDRLAGKPTHVNV 482
QY 478 SVVMAEVDGTC 488
Db 483 SVVMAXVXGPC 493

RESULT 15
US-09-833-245-2189
; Sequence 2189, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2189
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (488)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (490)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (494)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (495)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (505)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-2189

Query Match 83.7%; Score 2186; DB 11; Length 530;
Best Local Similarity 83.9%; Pred. No. 1.3e-121;
Matches 412; Conservative 21; Mismatches 50; Indels 8; Gaps 2;

QY 5 GFLWALVISICLFESMAQVOLVQSGAEVKKPGSSVRSVKASGCTPSSVAISWVROAPGQ 64
Db 4 GLTWFLVA-LLRGVHCQVQVSGAVVQPGGSLRLSCAASGFTFSRYGMHWVRQAPGK 62
QY 65 GLEWMGGIIPFGTANYAQKFGQGRVTITADESTSTAYMELSSLRSEDVAVYYCAR---- 119
Db 63 GLQWLALVLHDGQGYKYNEDVVKGRFTISRDNSSNNKYLQWDSLRLGEDTATYYCVRGMWEQ 122
QY 120 --DPFLHYWCGTLTVTVSTASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPOEPLSVTV 177
Db 123 LPSYYFDYMGQGTGLVTVSSASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPOEPLSVTV 182
QY 178 SESGQGVTAARNPPPSQDASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPC 237
Db 183 SESGQGVTAARNPPPSQDASGDLVTTSSQLTLPATQCLAGKSVTCHVKHYTNPSQDVTVPC 242
QY 238 PVSTPTPTSPSPPTPSPSCCHPRLSLHRPALEDLLLGSEANLCTCTLTGLRDASGVTF 297
Db 238 PVSTPTPTSPSPPTPSPSCCHPRLSLHRPALEDLLLGSEANLCTCTLTGLRDASGVTF 297
```

Search completed: October 25, 2005, 11:27:10  
Job time : 134.252 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:02:42 ; Search time 34.8456 Seconds  
(without alignments)  
1047.575 Million cell updates/sec

Title: US-10-644-256-3

Perfect score: 2611

Sequence: 1 MACPGFLWALVISTCLEFSM.....GKPHNVNVVMAEVDGTCY 489

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2186	83.7	530	4	US-09-800-729-112
2	2181	83.5	494	4	US-09-800-729-216
3	2053	78.6	499	3	US-09-049-672A-1
4	2004.5	76.8	483	3	US-09-049-672A-5
5	1970.5	75.5	487	4	US-09-800-729-145
6	1267.5	48.5	343	1	US-08-336-891-2
7	1267.5	48.5	343	5	PCT-US95-13795-4
8	880	33.7	476	2	US-08-378-939-10
9	853.5	32.7	462	4	US-09-627-896B-24
10	838.5	32.1	470	4	US-09-859-053-28
11	828.5	31.7	469	2	US-08-458-516-13
12	818	31.3	467	2	US-07-916-098A-45
13	817.5	31.3	445	3	US-08-341-560B-17
14	803	30.8	442	4	US-08-472-888A-7
15	803	30.8	442	5	PCT-US96-10043-9
16	801	30.7	442	1	US-08-480-036-2
17	801	30.7	442	1	US-08-461-968A-2
18	801	30.7	442	2	US-08-462-571-2
19	801	30.7	442	4	US-08-472-888A-2
20	801	30.7	442	5	PCT-US96-10043-12
21	800	30.6	442	1	US-08-461-968A-5
22	800	30.6	442	2	US-08-462-571-5
23	792.5	30.4	472	4	US-09-301-593-43
24	789	30.2	467	3	US-09-049-672A-8
25	767	29.4	552	5	PCT-US93-07832-23
26	765.5	29.3	473	4	US-09-828-995B-20
27	754.5	28.9	463	4	US-09-472-087-4

28	754.5	28.9	463	4	US-09-472-087-68	Sequence 68, Appl
29	754	28.9	468	3	US-09-485-737B-67	Sequence 67, Appl
30	754	28.9	469	4	US-10-071-485-67	Sequence 67, Appl
31	754	28.9	469	2	US-07-934-373C-23	Sequence 23, Appl
32	754	28.9	469	3	US-08-437-642B-23	Sequence 23, Appl
33	754	28.9	469	4	US-08-146-206C-23	Sequence 23, Appl
34	754	28.9	469	4	US-09-705-686-23	Sequence 23, Appl
35	754	28.9	469	4	US-09-705-392A-23	Sequence 23, Appl
36	754	28.9	469	4	US-09-705-398-23	Sequence 23, Appl
37	754	28.9	711	3	US-09-485-737B-90	Sequence 90, Appl
38	754	28.9	711	4	US-10-071-485-90	Sequence 90, Appl
39	753.5	28.9	468	2	US-08-116-247-7	Sequence 7, Appl
40	753.5	28.9	468	4	US-09-348-224-7	Sequence 7, Appl
41	753	28.8	450	2	US-08-788-800-12	Sequence 12, Appl
42	751	28.8	464	4	US-09-472-087-2	Sequence 2, Appl
43	751	28.8	464	4	US-09-472-087-56	Sequence 66, Appl
44	749	28.7	452	3	US-09-027-449-71	Sequence 71, Appl
45	749	28.7	452	3	US-09-026-985-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1  
US-09-800-729-112  
; Sequence 112, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 112  
; LENGTH: 530  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (488)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (490)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (494)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (495)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (505)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-800-729-112

Query Match	83.7%	Score 2186;	DB 4;	Length 530;
Best Local Similarity	83.9%	Pred. No. 9.7e-154;		
Matches 412;	Conservative 21;	Mismatches 50;	Indels 8;	Gaps 2;
Qy	5	GFLWALVISTCLEFSMAQVOLVQSGAEVVKKQSSVVRVSKASGGTFSSYALISWVQAPGQ	64	
Db	4	GLTWVFLVA-LLRGVHCQVQLVSGGAVVQPGGSLRLSCAASGFTFSRYGMHVRQAPGK	62	
Qy	65	GLEWNGGIPIFGTANYAAQKFGQRTVITADSTSTAYMELSSLRSEDTAVVYCAR----	119	
Db	63	GLQWALVHLDGGQKYNEDVVKGFTISRDNKNNKVLQMDSLRGEDTATVYCVGMWEQ	122	





```

; APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 145
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-800-729-145

Query Match 75.5%; Score 1970.5; DB 4; Length 487;
Best Local Similarity 76.0%; Pred. No. 7.7e-138;
Matches 381; Conservative 31; Mismatches 54; Indels 35; Gaps 6;

QY 7 LM-ALVISTCLFSPMAQVQLVQSGAEVKKPGSSVRVSCKASGTFSS--YAISVVRQAPG 63
DB 4 LWFELLVAAPRWLSQVQLQESGCLVKPSETLSLTCTVSGGSISGGHYWIRQHPG 63

QY 64 QGLEWGGIPIPTGANYAQKQGRVTTTADBSTSTAYWELSSURSEDYVYCARDFEL 123
DB 64 KGLEWI-GYISYNGVYTPNPSLKRVTISVDTSONQFSLRLSSVTAADYVYCAKD--- 119

QY 124 H-----YMGCGTLVTVSTASTPSPKVPFLSLCSTQPDGNVVIACLVQGGFF 168
DB 120 HRATRGYQLEYRGFDYWGQGLVTVSSASPTSPKVPFLSLDSTPQDGNVVIACLVQGGFF 179

QY 169 PQPFLSVTWSESGQVTAARNFPPPSODASGLYTTSSQLTLPATOCLAGKSVTCHVKYTN 228
DB 180 PQPFLSVTWSESGQVTAARNFPPPSODASGLYTTSSQLTLPATOCPDGKSVTCHVKYTN 239

QY 229 PSQDVTVPVSTPTPTSPSTPTPTSPSCCHPRLSLHRLPALEDLLLGSEANLCTLTGL 288
DB 240 PSQDVTVPVSTPTPTSPSTPTPTSPSCCHPRLSLHRLPALEDLLLGSEANLCTLTGL 286

QY 289 RDASGVFTFTWTPSSGKSAVOGPPDRDLCCGYSVSSVLSGCAEPNNHKGTFCTTAAYPESK 348
DB 287 RDASGATFTWTPSSGKSAVOGPPDRDLCCGYSVSSVLSGCAEPNNHKGTFCTTAAPHELK 346

QY 349 TPLTATLSKSGNTFRPEVHLLPPPEELALNELVTLTCLARGFSPKDVLRWLGSGQELP 408
DB 347 TPLTANITKSGNTFRPEVHLLPPPEELALNELVTLTCLARGFSPKDVLRWLGSGQELP 406

QY 409 REKYLTVASRQPSQGTTFVAVTSLRVAADWKKGDTFSCWVGHEALPLAFTQKTTIDRL 468
DB 407 REKYLTVASRQPSQGTTFVAVTSLRVAADWKKGDTFSCWVGHEALPLAFTQKTTIDRL 466

QY 469 AGKPTHVNVSVMAEVDGTCY 489
DB 467 AGKPTHVNVSVMAEVDGTCY 487

RESULT 6
US-08-336-891-2
; Sequence 2, Application US/08336891
; Patent No. 5622842
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN A
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 145
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-800-729-145

Query Match 48.5%; Score 1267.5; DB 1; Length 343;
Best Local Similarity 67.0%; Pred. No. 5.1e-86;
Matches 236; Conservative 38; Mismatches 69; Indels 9; Gaps 2;

QY 138 SPTSPPKVPFLSLCSTQPDGNVVIACLVQGGFFPQBPPLSVTWSESGQVTAARNFPPSQDASG 197
DB 1 SKTSPSVFPLSLCHQSEGEYVIGCLVQGGFFPPEPVNVTWNAKDS*SVKNFPPMKAATG 60

QY 198 DLYTTSSQLTLPATOCLAGKSVTCHVKYTNPSQDVTVPVSTPTPTPTPTPTPTPTPTPTPT 257
DB 61 SLYTSSQLTLPAAQCPDDSSVKQVQHASPSKAVSVP-----KNSHPCHPCPS 112

QY 258 CCHPRLSLHRLPALEDLLLGSEANLCTLTGLRDASGVFTWTPSSGKSAVOGPPDRDLQ 317
DB 113 CNEPRLSLQKPALEDLLLGSEANLCTLTGLKDPKGAFTFWNPSKGKEPIQKPERDSCG 172

QY 318 CYSVSVLSGCAEPNNHKGTFCTTAAYPESKTPLTATLSKSGNTFRPEVHLLPPPSBELA 377
DB 173 CYSVSVLPVGCADPNHGDFTCTATHPESKSPITVSTTKTEHIPQVHLLPPPSBELA 232

QY 378 LNELVTLCLARGFSPKDVLRWLGSGQELPPELTLTCLARGFSPKDVLRWLGSGQELP 437
DB 233 LNELVTLCLVGRGFKDVLVRLQGTQELPQEKYLTWEPLKEPDQ-TNMFVATSMILRVT 291

QY 438 ASDWKKGDTFSCWVGHEALPLAFTQKTTIDRLAGKPTHVNVSVMAEVDGTCY 489
DB 292 AEDWQGEKFSWVGHEALPLMSFTQKTIIDRLAGKPTHVNVSVMAEVDGTCY 343

RESULT 7
PCT-US95-13795-4
; Sequence 4, Application PC/TUS9513795
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13795
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

```

```
; NAME: CARTY, CHRISTINE E.
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19211Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-6734
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-13795-4

Query Match      48.5%; Score 1267.5; DB 5; Length 343;
Best Local Similarity 67.0%; Pred. No. 5.1e-86;
Matches 236; Conservative 38; Mismatches 69; Indels 9; Gaps 2;

Qy 138 SPTSPKVPPLSLCTQPDGNVVIACLVQGFPPQBPPLSVTHWSESGQGVTAARNFPPSQDASG 197
Db 1 SKTSPSPVPLSLCHQSESEGVVIGCLVQGFPPPEPVNVNTWNAKGDSTSVKNFPPMKAATG 60

Qy 198 DLYTTSSQLTLPATQCLAGKSVTCHVHYTNPSODVTPCPVPSTPTPTSPSTPTSPS 257
Db 61 SLYTMSQLTLPAAQCPDDSVKQVQHASPSKAVSVPC-----KNSHPCHECPSS 112

Qy 258 CCHPRLSLHRLPALEDLLGSEANLTCTLTGLDASGVTFWTWPSGSKSAVQGPDRDLG 317
Db 113 CNEPRLSLQRLPALEDLLGSEANLTCTLTGLKOPKGAFTWNPSPGKEPIQKNPERDSG 172

Qy 318 CYSVSVSLGCAEPWNGHKTFTCTAAVPESKTPLTATLSKSGNTFRREVHLLPPPSBELA 377
Db 173 CYSVSVSLPGCADPWNHGDFTSCATHPESKSPITVSITKTEIHPQVHLLPPPSBELA 232

Qy 378 LNELVTLTCLARGSPKDLVRLVQSGOELPREKYLTVASRQEPSQGTTFVAVTSILRVA 437
Db 233 LNELVTLTCLVRGPKDLVRLVQSGOELPREKYLTVASRQEPSQGTTFVAVTSILRVA 291

Qy 438 AEDWKQKGDFTSCMGHEALPLAFTQKTIDRLAGKPTHVNSVVAEVDGTCY 489
Db 292 AEDWKQKGDFTSCMGHEALPLAFTQKTIDRLAGKPTHVNSVVAEVDGTCY 343

RESULT 8
US-08-378-939-10
; Sequence 10, Application US/08378939
; Patent No. 5876961
; GENERAL INFORMATION:
; APPLICANT: CROME, JAMES SCOTT
; APPLICANT: LEWIS, ALAN PETER
; TITLE OF INVENTION: PRODUCTION OF ANTIBODIES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 THIRTEENTH ST. N.W.
; CITY: WASHINGTON
; STATE: D. C.
; COUNTRY: U. S.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/378,939
; APPLICATION NUMBER: US/08/378,939
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952640
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: ERNST, BARBARA G
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1808-118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 783-6040
; TELEFAX: (202) 783-6031
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-939-10

Query Match      33.7%; Score 880; DB 2; Length 476;
Best Local Similarity 42.1%; Pred. No. 3.6e-57;
Matches 209; Conservative 69; Mismatches 161; Indels 58; Gaps 17;

Qy 6 FLNALVISTCLEFSMAQVLQSGAEVKKPGSSVRVSKASGGTFSSYATISWVRQAQGG 65
Db 7 FLFWAAATGVQ---SQMVVQSGAEVKKPGSSVTVSKASGGTFSSYATISWVRQAQGG 63

Qy 66 LEWGGIIPFGTANTYAQKFGQVRVITADSTSTAYMELSSLSRSEDVAVYYCARDPF--- 122
Db 64 LEWGGIIPFGTPTYSQNFQGRVITADKSTSTAHMELTSLRSEDVAVYYCATDVRQA 123

Qy 123 -----LHYVGGTGLTVTVSTASPTSPKVPFLSLCS--TOPDGNVVIACLVQGFPPQEP 172
Db 124 NFDRAVRVGFDPWQGGTLTVTVSSASTKGPSVFFPLAPSSKSTSGTAAALCLVKDYFP-EP 182

Qy 173 LSVTWSESGOGVTAARNFPPSQDASGLYTTSSQLTLPATQCLAGKSVTCHVHY-TNPSQ 231
Db 183 VTVSWNSGALTGVHFTFAVLQSSG-LYSLSVVTVPSSS-LGTQTICNVNHPSPNTKV 240

Qy 232 DVTV-----PCVPSTPTPTSPST-PTPTSPSCCHPRLSLHRLPALEDLLGSE 278
Db 241 DKKVEPKSCDKHTCTCPAPPELLGGPSVFLFPPKPK-----DTLMISRT 285

Qy 279 ANLTCTLTGL-RDASGVTFWTWPSGK--SAVQGPDRDLGCGYSVSVSLGCAEPWNGH 335
Db 286 PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG 345

Qy 336 KTFCTAAVPESKTPLTATLSKS-GNTFRREVHLLPPPSBELALNELVTLTCLARGSPK 394
Db 346 KEYCKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQ-VSLTCLVKGYEFS 404

Qy 395 DVLVRVWLGSGOELPREKYLTVASRQEPSQGTTFVAVTSILRVAEDNKKGDFTSCMGVHE 454
Db 405 DIAVEVESNQ--PENNYKTTTPPVLD-SDG--SFPLYSKLTVDKSRNQGNVFCVSWH 459

Qy 455 ALPLAFTQKTIDRLAGK 471
Db 460 ALHNHYTQKSLSLSPGK 476
```

```
RESULT 9
US-09-627-896B-24
; Sequence 24, Application US/09627896B
; Patent No. 6827934
; GENERAL INFORMATION:
; APPLICANT: CO, MAN SUNG
; APPLICANT: VASQUEZ, MAXIMILIANO
; APPLICANT: CARRENO, BEATRIZ
; APPLICANT: CELNIKER, ABBIE CHERYL
; APPLICANT: COLLINS, MARY
; APPLICANT: GOLDMAN, SAMUEL
; APPLICANT: GRAY, GARY S.
; APPLICANT: KNIGHT, ANDREA
; APPLICANT: O'HARA, DENISE
; APPLICANT: RUP, BONITA
; APPLICANT: VELDMAN, GEERTRUIDA M.
; TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH B7-2 AND METHODS
; TITLE OF INVENTION: OF TREATMENT THEREWITH
```

```
/ FILE REFERENCE: 08702.0081-01000
/ CURRENT APPLICATION NUMBER: US/09/627,896B
/ CURRENT FILING DATE: 2000-07-27
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: Patent in Ver. 2.1
/ SEQ ID NO 24
/ LENGTH: 462
/ TYPE: PRT
/ ORGANISM: 3D1 heavy chain
US-09-627-896B-24

Query Match      32.7%; Score 853.5; DB 4; Length 462;
Best Local Similarity 41.9%; Pred. No. 3.2e-55;
Matches 196; Conservative 70; Mismatches 161; Indels 41; Gaps 15;

QY 21 AQVQLVQSGAEVKKPGSSVRVSKASGGTFSSYAIASWVRQAPQGLQEMWGGIPIGTAN 80
DB 19 SQVQLVQSGAEVKKPGSSVRVSKASGGTFSSYAIASWVRQAPQGLQEMWGGIPIGTAN 78
QY 81 YAKQFGQGVTVITADESTSTAYMELSLRSRSDTAVYYCARDP-FLHYWGQGLTVVSTASP 139
DB 79 YNKKFKGKATMTVDKSTSTAYMELSLRSRSDTAVYYCARAAWYMDYWGQGLTVVSSAST 138
QY 140 TSPKVPFLSLCS--TOPDGNVIAACLVQGFPPQBPPLSVTWSGQGVYARNFPSPQASGD 198
DB 139 KGPSVFLPAPCSRSTSESTAALGCLVKDYPP-EPVTVSNWSGALTSGVHTFPAVLQSSG- 196
QY 199 LYTTSSQLTLPATQCLAGKSVTKVHKYTNPSQDV-----TVPCVPSPSTPTSPST 250
DB 197 LYSLSVTVTPSSN-FGTQTYTCNVDPKPSNTKVDKTKVERKCCVCPCCPAPAAAPSV 255
QY 251 ---PPTPSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGL-RDASGVTVFTWTPSSGK-- 304
DB 256 FLFPKPK-----DTLMISRTPEVTCVVVDVSHEDDEVQFNWYVDGVEVH 300
QY 305 SAVQGPDPDRDLCCGYSSVSLSCAEPNHNKGTFTCTTAAYPESKTPLTATLSKS-GNTER 363
DB 301 NAKTKPREEQFNSTFRVSVLTVVHQDLNKGKEYKCKVSKNKGKLPAPIETISKTKQPRE 360
QY 364 PEVHLPPPPSEELALNELVLTCLAGFSPKDVLRWLGQSQELPREKYLTVASRQEPSPQ 423
DB 361 PQYITLPPSREEMTKNQ-VSLTCLVKGFPSPDI AVEWESNGQ--PENNYKTPPMLD-SD 416
QY 424 GTTTFVAVTSLRVAEDWKKGDTFSCMVGHEALPLAFTOKTIDRLAGK 471
DB 417 G--SFFLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSLSPGK 462

RESULT 10
US-09-859-053-28
/ Sequence 28, Application US/09859053
/ Patent No. 6803039
/ GENERAL INFORMATION:
/ APPLICANT: Tezuka, Takashi
/ APPLICANT: Hori, Noriaki, Takashiro
/ TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
/ TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE AILIM AND
/ TITLE OF INVENTION: PHARMACEUTICAL USE THEREOF
/ FILE REFERENCE: 06501-079001
/ CURRENT APPLICATION NUMBER: US/09/859,053
/ CURRENT FILING DATE: 2001-05-16
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR FILING DATE: 2000-05-18
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 28
/ LENGTH: 470
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-859-053-28
```

```
Query Match      32.1%; Score 838.5; DB 4; Length 470;
Best Local Similarity 41.7%; Pred. No. 4.2e-54;
Matches 200; Conservative 64; Mismatches 159; Indels 57; Gaps 18;

QY 21 AQVQLVQSGAEVKKPGSSVRVSKASGGTFSSYAIASWVRQAPQGLQEMWGGIPIGTAN 80
DB 19 SQVQLVQSGAEVKKPGASVKASKASGYFTFTGYMHWVRQAPQGLQEMWGWINPHSGGTN 78
QY 81 YAKQFGQGVTVITADESTSTAYMELSLRSRSDTAVYYCAR-----DPLHYWGQ 129
DB 79 YAKQFGQGVTVITRDTSTISYAMELSLRSRSDTAVYYCARTYYDSSGGYVHDAP-DIWQGG 137
QY 130 TLVTVSTASPTSPKVPFLSLCS--TOPDGNVIAACLVQGFPPQBPPLSVTWSGQGVYARN 188
DB 138 TWTVSSASTKGPSVFLPAPCSRSTSESTAALGCLVKDYPP-EPVTVSNWSGALTSGVHT 196
QY 189 FPPSQDASGLTYTSSQLTLPATQCLAGKSVTKVHKY-TNPSQDVTV-----PCP 238
DB 197 FPAVLQSSG-LYSLSVTVTPSSN-FGTQTYTCNVDPKPSNTKVDKTKVERKCCVCPCCP 254
QY 239 VSTPTPTSPST---PPTPSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGL-RDASGV 294
DB 255 ---APPVAGPSVFLFPKPK-----DTLMISRTPEVTCVVVDVSHEDDEV 296
QY 295 TPTWTPSSGK--SAVQGPDPDRDLCCGYSSVSLSCAEPNHNKGTFTCTTAAYPESKTPLT 352
DB 297 QFNWYVDGVEVHNKTKPREEQFNSTFRVSVLTVVHQDLNKGKEYKCKVSKNKGKLPAP 356
QY 353 ATLSKS-GNTPFPEVHLPPPPSEELALNELVLTCLAGFSPKDVLRWLGQSQELPREK 411
DB 357 KTIKTKGQPREPQVYTLPPSREEMTKNQ-VSLTCLVKGFPSPDI AVEWESNGQ--PENN 413
QY 412 YLTWASRQSPSGTTFVAVTSLRVAEDWKKGDTFSCMVGHEALPLAFTOKTIDRLAGK 471
DB 414 YKTTTPMLD-SDG--SFFLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSLSPGK 470

RESULT 11
US-08-458-516-13
/ Sequence 13, Application US/08458516
/ Patent No. 5777085
/ GENERAL INFORMATION:
/ APPLICANT: Co, Man Sung
/ APPLICANT: Tso, J. Yun
/ TITLE OF INVENTION: Humanized Antibodies Reactive with
/ TITLE OF INVENTION: GPIIb/IIIa
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: William M. Smith
/ STREET: One Market Plaza, Steuart Tower, Suite 2000
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94105
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/458,516
/ FILING DATE:
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/059,159
/ FILING DATE: 03-MAY-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Smith, William M.
/ REGISTRATION NUMBER: 30,223
/ REFERENCE/DOCKET NUMBER: 11823-37-3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-326-2400
```

```
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-458-516-13

Query Match      31.7%; Score 828.5; DB 1; Length 449;
Best Local Similarity 41.9%; Pred. No. 2.2e-53;
Matches 198; Conservative 64; Mismatches 164; Indels 47; Gaps 16;

Qy 22 QVQLVQSGAEVKKPGSSVRVSCKASGCTFSSYAIISWRQAPGQGLEWMGGIPIFGTANY 81
Db 1 QVQLVQSGAEVKKPGSSVRVSCKASGCTFSSYAIISWRQAPGQGLEWMGGIPIFGTANY 60
Qy 82 AQKQGRVTITADESTAYMELSSLRSEDTAVVYCARDP-----FLHYWGQGLTVTSTA 137
Db 61 NEKFKGRVTLTVDESTNTAYMELSSLRSEDTAVVYCARDP-----FLHYWGQGLTVTSTA 120
Qy 138 SPTSPKVPFLPLSCS--TPDGNVITACLVQGFPPORPLSVTWSESGQGVTAARNPPSQDAS 196
Db 121 STKGPSVPLAPSSKSTSGGTAALGCLVKDYFP-EPVTVSNWNGALTSGVHTFPVAVLQSS 179
Qy 197 GDLTYTSSQLTLPATCLAGKSVTCHVKHY--TNPSQDVTV-----PCVPSTP 243
Db 180 G-LYSLSSVVTPSS--LGTQYICNVNHPKSNKVDKKVPEKSCDKHTHTCPCPAPPELL 237
Qy 244 PTPSPST--PTTPSPSCCHPRLSLRPALEDLLLGSEANLTCITLGL-RLDASGVTFPTWPS 301
Db 238 GGPVSFVLPFPKPK-----DTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 282
Qy 302 SGK--SAVQGPDRDLCCGYSSVSLSCAPRNHNGKFTCTTAAYPESKTPLTATLTKS- 358
Db 283 GVEVHNAKTKPREQVNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPAPIETKISRAK 342
Qy 359 GNTFRPEVHLLPPPSSEELALNELVTLTCLARGSPKQVLRWLGSGQBELPREKYLTVASR 418
Db 343 GQPREPQVYTLPPSRDELTKNQ-VSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTTPV 399
Qy 419 QPSQGTITFAVTSILRVAEDWKKGTDFSCVNGHEALPLAFTOKTIDRLAGK 471
Db 400 LD-SDG--SFFLYSKLTVDKSRWQGNVFCGSMVHEALHNHYTQKSLSLSPGK 449

RESULT 12
US-07-916-098A-45
; Sequence 45, Application US/07916098A
; Patent No. 5871732
; GENERAL INFORMATION:
; APPLICANT: BURKLY, LINDA C.
; APPLICANT: CHISHOLM, PATRICIA L.
; APPLICANT: THOMAS, DAVID W.
; APPLICANT: ROSA, MARGARET D.
; APPLICANT: ROSA, JOSEPH J.
; TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
; TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
; STREET: 10 SOUTH WACKER DRIVE
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/07/916,098A
; FILING DATE: July 24, 1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08843
; FILING DATE: No. 5871732ember 27, 1991
; CLASSIFICATION: 424
; APPLICATION NUMBER: 07/618,542
; FILING DATE: No. 5871732ember 27, 1990
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: JOHN J. MC DONNELL
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,310-G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 715-1000
; TELEFAX: (312) 715-1234
; TELEX: 910/221-5317
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-916-098A-45

Query Match      31.3%; Score 818; DB 2; Length 467;
Best Local Similarity 40.8%; Pred. No. 1.4e-52;
Matches 194; Conservative 71; Mismatches 159; Indels 52; Gaps 17;

Qy 21 AQQLVQSGAEVKKPGSSVRVSCKASGCTFSSYAIISWRQAPGQGLEWMGGIPIFGTAN 80
Db 19 SQVQLQSGAEVKKPGASVKVSCKASGYTFTSYVIHWVRQAPGQGLEWIGVINPYNDGTD 78
Qy 81 YAKQFGRVTITADESTAYMELSSLRSEDTAVVYCARD-----PFLHYWGQGLTVT 133
Db 79 YDEKFKRATVTLDPSTNTAYMELSSLRSEDTAVVYCARDKDNVATGANFAYWQGLTVT 138
Qy 134 VSTASPTSPKVPFLPLSCS--TPDGNVITACLVQGFPPORPLSVTWSESGQGVTAARNPPSP 192
Db 139 VSSASTKGPSVFLPLAPCSRSTSESTAALGCLVKDYFP-EPVTVSNWNGALTSGVHTFPV 197
Qy 193 QDASGDIYTTSSQLTLPATCLAGKSVTCHVKHY--TNPSQDVTV-----PCVPSTPPT 245
Db 198 LQSSG-LYSLSSVVTPSS--LGTQYTCNVNHPKSNKVDKRVESKYGPCCP---SCPA 252
Qy 246 P-----SPST---PPTSPSCCHPRLSLRPALEDLLLGSEANLTCITLGL-RLDASGVTF 297
Db 253 PEFLLGGPSVFLFPKPK-----DTLMISRTPEVTCVVVDVSHEDPEVQFN 297
Qy 298 WTPSSGK--SAVQGPDRDLCCGYSSVSLSCAPRNHNGKFTCTTAAYPESKTPLTATL 355
Db 298 WYVDGVEVHNAKTKPREQVNSTYRVSVLTVLHQDLNKGKEYCKVSNKGLPSSIEKTI 357
Qy 356 SKGNTFRPEVHLLPPPSSEELALNELVTLTCLARGSPKQVLRWLGSGQBELPREKYLTV 415
Db 358 SKAQPREPQVYTLPPSQEEMTKNQ-VSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTT 414
Qy 416 ASRQEPSQGTITFAVTSILRVAEDWKKGTDFSCVNGHEALPLAFTOKTIDRLAGK 471
Db 415 PPVLD-SDG--SFFLYSKLTVDKSRWQGNVFCGSMVHEALHNHYTQKSLSLSPGK 467

RESULT 13
US-08-341-560B-17
; Sequence 17, Application US/08341560B
; Patent No. 6165745
; GENERAL INFORMATION:
; APPLICANT: Ward, E. Sally
; APPLICANT: Kim, Jin-Kyoo
; TITLE OF INVENTION: Recombinant Production of
; TITLE OF INVENTION: Immunoglobulin-Like Domains in Prokaryotic Cells
; NUMBER OF SEQUENCES: 33
```



Db 239 PPKPK-----DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 283  
Qy 308 QGPPDRDLCCGYSSVSSVLSCAEPWNHKGKTCTCTAAAPESKTPLTATLSKS-GNTFRPEV 366  
Db 284 TKPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKVSNKALPAPIEKTISKAKGQPREPQV 343  
Qy 367 HLLPPPEELALNELVLTCLARGFSPKDVLRWLQSQELPREKYLTVASRQPSOGTT 426  
Db 344 YTLPPSRDELTKNQ-VSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTTPVLD-SDG-- 397  
Qy 427 TFAVTSILRVAEADWKGDTSFSCMVGHEALPLAFTQKTIIDLACK 471  
Db 398 SFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 442

RESULT 15  
PCT-US96-10043-9  
; Sequence 9, Application PC/TUS9610043  
; GENERAL INFORMATION:  
; APPLICANT: The General Hospital Corporation  
; TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES  
; TITLE OF INVENTION: AND METHODS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10043  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/000,213  
; FILING DATE: 14-JUN-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lech, Karen F.  
; REGISTRATION NUMBER:  
; REFERENCE/DOCKET NUMBER: 00786/284001  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 442 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US96-10043-9

Query Match 30.8%; Score 803; DB 5; Length 442;  
Best Local Similarity 42.4%; Pred. No. 1.7e-51;  
Matches 197; Conservative 57; Mismatches 151; Indels 60; Gaps 18;  
Qy 21 AQVLVQSGAEVKKPGSSVRVSCKASGCTFSYSAISWVRQAPGQGLEWMGGIIPFGTAN 80  
Db 24 SQVLVQSGAEVKKPGSSVRVSCKASGCTFSYSAISWVRQAPGQGLEWMGGIIPFGTAN 83  
Qy 81 YAKQFGQGVTTITADESTAYMELSLRSEDTAVYYCAR-----DPFLHYWGQGLV 132  
Db 84 YAKQFGQGVTTITADEST-----ARDNGA-YCSGGSGYSGWFDP-----WGQGLV 127

Qy 133 TVSTASTSPKVPPLSLCS-TQPDGNVVVIACLVQGFPPQBPPLSVTWSESGQGVTAARNFPP 191  
Db 128 TVSSASTKGPSPVFLAPSSKSTSGTAAALGCLVKDYFP-BFVTVSMNSGALTSGVHTFPA 186  
Qy 192 SQDASGDLVTTSSQLTLPATQCLAGKSVTCHVXHYTNPSQDVTVPQVPSPTPTPSPST- 250  
Db 187 VLOSSG-LYSLSSVVTVPSDDKKVEPKSCD-KTHTCP-----PCPAPELLGGPSVFLF 238  
Qy 251 PPTPSPSCCHPRLSLHRPALELLLGSEANLTCTLTGL-RDASGVTTTWTTPSSCK--SAV 307  
Db 239 PPKPK-----DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 283  
Qy 308 QGPPDRDLCCGYSSVSSVLSCAEPWNHKGKTCTCTAAAPESKTPLTATLSKS-GNTFRPEV 366  
Db 284 TKPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKVSNKALPAPIEKTISKAKGQPREPQV 343  
Qy 367 HLLPPPEELALNELVLTCLARGFSPKDVLRWLQSQELPREKYLTVASRQPSOGTT 426  
Db 344 YTLPPSRDELTKNQ-VSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTTPVLD-SDG-- 397  
Qy 427 TFAVTSILRVAEADWKGDTSFSCMVGHEALPLAFTQKTIIDLACK 471  
Db 398 SFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 442

Search completed: October 25, 2005, 11:09:31  
Job time : 36.8456 secs

**THIS PAGE BLANK (USPIC)**